Author:
Biggs, Kirsti D

Title:
On additive problems involving shifted integers and ellipsephic sets

ON ADDITIVE PROBLEMS INVOLVING SHIFTED INTEGERS AND ELLIPSEPHIC SETS

Kirsti Debrah Biggs

A dissertation submitted to the University of Bristol in accordance with the requirements for award of the degree of Doctor of Philosophy in the Faculty of Science

School of Mathematics, August 2019

Abstract

In this thesis, we present a series of results concerning the number of solutions to equations and inequalities involving sums of perfect powers. In order to count such solutions, we use variants of the Hardy–Littlewood circle method, a versatile technique whose history is outlined in Chapter 1.

Firstly, to handle inequalities we use the Davenport–Heilbronn version of the circle method, in the form developed by Freeman, which we provide an introduction to in Chapter 2. In Chapter 3, we apply this method to the problem of counting solutions to an inequality in which our variables have been shifted by small real numbers. Specifically, for natural numbers k and s and real numbers $\theta_1, \dots, \theta_s \in (0,1)$ with θ_1 irrational, let $N(\tau)$ be the number of solutions in positive integers x_i to the inequality

$$\left|(x_1 - \theta_1)^k + \dots + (x_s - \theta_s)^k - \tau\right| < 1.$$

We show that an asymptotic formula for $N(\tau)$ holds whenever $k \geq 4$ and $s \geq k^2 + (3k-1)/4$, an improvement on a result of Chow. We also prove a shifted analogue of a result of Wright showing that such an inequality does not always have solutions in which the variables are forced to lie in a short interval.

We then turn to problems involving integers whose digits in a given base are restricted to certain values; we refer to such integers as *ellipsephic*. We use Wooley's efficient congruencing method to bound the number of ellipsephic solutions to the Vinogradov system

$$x_1^j + \dots + x_s^j = y_1^j + \dots + y_s^j, \quad (1 \leq j \leq k),$$

handling the case $k = 2$ in Chapter 4, and the general case in Chapter 5. The additive structure of our ellipsephic sets enables us to achieve better bounds than those previously available, since any application of earlier results would use only the density of the set of variables within the natural numbers. For example, in the case where our variables have square digits, we obtain diagonal behaviour with twice as many variables as in the classical case.

Acknowledgements

I would like to thank my supervisor, Trevor Wooley, for taking me on as a student, and for his dedicated supervision and continued support and guidance during my PhD. His numerous ideas and suggestions and our many mathematical discussions have been invaluable throughout this process.

I would also like to thank the mathematical community at the University of Bristol, including, but not limited to, my fellow PhD students and the regular crowd at the Vic, as well as my contacts and collaborators further afield.

Finally, thank you to my family, who have always been supportive of my mathematical ambitions, and especially to my parents, Debrah and Geoff, who taught me to count at an early age. I don't expect you to read the entirety of my thesis, but it all boils down to counting in the end.

Author's Declaration

I declare that the work in this dissertation was carried out in accordance with the requirements of the University's Regulations and Code of Practice for Research Degree Programmes and that it has not been submitted for any other academic award. Except where indicated by specific reference in the text, the work is the candidate's own work. Work done in collaboration with, or with the assistance of, others, is indicated as such. Any views expressed in the dissertation are those of the author.

SIGNED: ..... DATE:.....

Contents

'Arithmancy looks terrible,' said Harry, picking up a very complicated-looking number chart.

'Oh, no, it's wonderful!' said Hermione earnestly. 'It's my favourite subject!'

—*J. K. Rowling, Harry Potter and the Prisoner of Azkaban*

Chapter 1

Introduction

1.1 Background and motivation

For millennia, Diophantine problems—that is to say, problems concerning integer solutions to equations with integer coefficients—have occupied the minds of many mathematicians. Notable examples include the Fermat equation $x^n + y^n = z^n$, famously proved to have no non-trivial integer solutions when $n \geq 3$ by Wiles in [63], and Pell's equation $x^2 - ny^2 = 1$, which was studied extensively by Brahmagupta in the 7th century (see, for example, [24, Chapter XII] for an account of the history of this problem). Given such an equation, or system of equations, we wish to determine whether solutions exist, and if so, we wish to count them or estimate their number, and to investigate the structure of the set of solutions.

In 1770, Waring conjectured in [61] that every natural number can be written as a sum of at most nine positive cubes, at most nineteen fourth powers, 'and so on'. The corresponding result in the quadratic case, that every natural number can be written as a sum of at most four squares, was proved by Lagrange in the same year. In its original form, Waring's conjecture remained open until 1909, when Hilbert proved in [35] that for any natural number $k \geq 2$, there exists $s = s(k)$ such that any natural number n can be written as

$$n = x_1^k + \ldots + x_s^k, \tag{1.1.1}$$

where $x_i \in \mathbb{N}_0 = \mathbb{N} \cup \{0\}$ for $1 \leq i \leq s$. The history of this problem and progress towards it until 2002 is covered comprehensively in the survey paper

[57] of Vaughan and Wooley, which we have drawn upon in preparing this section.

The value of the minimal such s is now essentially known. As is usual in this field, let $g(k)$ be the smallest value of s for which any $n \in \mathbb{N}$ can be written in the form (1.1.1). Write $\lfloor x \rfloor$ for the floor function of x and $\{x\}$ for the fractional part $x - \lfloor x \rfloor$. Then

$$g(k) = \begin{cases} 2^k + \lfloor (3/2)^k \rfloor - 2, & \text{if } 2^k\{(3/2)^k\} + \lfloor (3/2)^k \rfloor \leq 2^k, \\ 2^k + \lfloor (3/2)^k \rfloor + \lfloor (4/3)^k \rfloor - \theta, & \text{otherwise,} \end{cases}$$

where

$$\theta = \begin{cases} 2, & \text{if } \lfloor (4/3)^k \rfloor \lfloor (3/2)^k \rfloor + \lfloor (4/3)^k \rfloor + \lfloor (3/2)^k \rfloor = 2^k, \\ 3, & \text{if } \lfloor (4/3)^k \rfloor \lfloor (3/2)^k \rfloor + \lfloor (4/3)^k \rfloor + \lfloor (3/2)^k \rfloor > 2^k. \end{cases}$$

It is known by a result of Mahler from 1957, in [45], that the first case, namely the case where $g(k) = 2^k + \lfloor (3/2)^k \rfloor - 2$, holds for all but at most a finite number of values of k. However, even under the assumption that this case occurs for all k, the value of $g(k)$ obtained is rather large and is heavily skewed by the number of kth powers required to represent small values of n.

As such, there is significant interest in the related question of the minimal s with the property that any *sufficiently large* value of n may be written in the form (1.1.1), commonly denoted by $G(k)$. In this case, the exact value of such a minimal s is known only for $k = 2$ and 4, with $G(2) = 4$ due to Lagrange (see [40]), and $G(4) = 16$ to Davenport in [20]. For the case $k = 3$, Linnik proved in [43] that $G(3) \leq 7$, but this is the best bound to date; it is conjectured that $G(3) = 4$. Numerically speaking, Dickson proved in [23] that all integers except 23 and 239 are sums of eight cubes, and Siksek proved in [53] that all integers greater than 454 are sums of at most seven cubes. For large k, the best upper bounds are due to Wooley (see [64]) and take the form $G(k) \leq (1 + o(1))k \log k$.

Further, if solutions to (1.1.1) exist, we are interested in determining their number, so we may also consider the minimal s for which an asymptotic formula for the number of such solutions holds, which we denote by $\tilde{G}(k)$. The resolution of the main conjecture in Vinogradov's mean value theorem (see later in this section) allowed Bourgain to prove in [10] that $\tilde{G}(k) \leq$

$k^2 - k + O(\sqrt{k})$, and this was further refined by Wooley in [71] to give $\tilde{G}(k) \leq k^2 - k + 2\lfloor\sqrt{2k+2}\rfloor - \theta(k)$, where $\theta(k)$ is 1 or 2 depending on an explicit condition on k.

The techniques used to obtain many of the above bounds, as well as numerous other results on additive problems, go by the name of the *(Hardy–Littlewood) circle method*, for historical reasons which we outline now. In [32], Hardy and Ramanujan proved an asymptotic formula for the number of partitions of a natural number, using a dissection of the unit circle into arcs which appears to have been the origin of the circle method. They went on to mention potential applications to two problems concerning sums of squares: counting partitions of a number into squares, and counting representations of a number as a sum of a predetermined number of squares—this latter case being the one of interest to us in the context of Waring's problem.

In [29] and [30], Hardy and Littlewood developed this further to become the method which now bears their name, and presented their new solution to Waring's problem. In the former paper, they observed that, unlike the method of Hilbert (as simplified by Remak in [51]):

> "This solution is not, in any sense of the word, elementary. It is based throughout on Cauchy's theorem and the ordinary machinery of the theory of analytic functions, and has, from beginning to end, no point of contact with Hilbert's solution. It might seem that a highly transcendental proof ... is unnecessary. This view, we think, would rest upon a misapprehension. It seems to us most desirable and important that Waring's Problem ... should be brought into relation with the transcendental side of the Analytic Theory of Numbers. Further, the method which we follow ... is a method of great power and wide scope, applicable to almost any problem concerning the decomposition of integers into parts of a particular kind".

Hardy and Littlewood's assessment of their method as powerful and widely applicable proved accurate; since its development, the method has been brought to bear on a multitude of related problems in the field of Diophantine equations. One such concerns the representation of integers as sums of primes. Vinogradov proved that all sufficiently large odd numbers are the sum of three primes (see [59, Chapter X]), and Helfgott extended this result, supported

by computation of Helfgott and Platt (see [33] and [34]), to all odd numbers greater than five, thereby settling the so-called *ternary Goldbach conjecture*. The corresponding conjecture that all even numbers greater than two are the sum of two primes remains open. The circle method has also been used in Birch's theorem on simultaneous solutions to rational forms in many variables (see [8]), and Roth's theorem on sets of natural numbers lacking three-term arithmetic progressions (see [52]). Furthermore, there is work on numerous versions of Waring's problem in other settings, such as number fields and function fields, and with restricted sets of variables, such as the Waring–Goldbach problem, which, as the name might suggest, requires the variables used in Waring's problem to be prime. For further details on the Hardy–Littlewood circle method beyond those presented here, see the book [21] by Davenport or the book [56] by Vaughan.

The modern version of the method owes much to Vinogradov, who, in [58], simplified the work of Hardy and Littlewood into something resembling its current form, replacing their contour integrals and power series with finite exponential sums via the key observation, on which the circle method as we know it today hinges, that we have

$$\oint e(\alpha n)\,d\alpha = \begin{cases} 1, & \text{if } n = 0, \\ 0, & \text{if } n \in \mathbb{Z} \setminus \{0\}, \end{cases} \tag{1.1.2}$$

where we write $e(z) = e^{2\pi i z}$, and $\oint$ denotes here the integral over the unit interval $[0,1]$, and in general the integral over the unit cube $[0,1]^k$ of appropriate dimension. Consequently, we may replace n in (1.1.2) by any expression taking integral values in order to determine whether that expression is zero, and therefore whether a given choice of variables solves an equation. In order to count the number of solutions, we sum over the potential values of the variables. For example, if we wish to know the number of integer solutions to the equation $g(x) = 0$, we evaluate

$$\oint \sum_{x \in \mathbb{Z}} e(\alpha g(x))\,d\alpha,$$

and if we are only interested in the existence of solutions, it suffices to bound the above expression away from zero.

In order to evaluate or bound the relevant integral, we split the domain $[0,1]$ into two parts, known as the major and minor arcs. The major arcs are small intervals around rational numbers with small denominators, where we expect that the integrand may be large, and the minor arcs are the remainder, where we expect significant cancellation within the sum. As such, the major arcs should provide the main term in any expression for the number of solutions, and the minor arcs contribute only to the error.

In the specific case of Waring's problem, we write $P = n^{1/k}$ and

$$f(\alpha) = \sum_{1 \leq x \leq P} e(\alpha x^k),$$

and consider the integral

$$R_{s,k}(n) := \oint f(\alpha)^s e(-\alpha n)\, d\alpha,$$

which counts the number of ways to write n as a sum of s kth powers.

Typically, around a point a/q with $a \in \mathbb{Z}$ and $q \in \mathbb{N}$ coprime, we define the major arc

$$\mathfrak{M}(q,a) = \{\alpha \in [0,1] \mid |\alpha - a/q| \leq QP^{-k}\},$$

where $Q = P^\delta$ is some small power of P. On this arc, we expect

$$f(\alpha) \sim q^{-1} S(q,a) \nu(\alpha - a/q),$$

where

$$S(q,a) = \sum_{r=1}^{q} e(ar^k/q)$$

and

$$\nu(\beta) = \int_0^P e(\beta\gamma^k)\, d\gamma.$$

This asymptotic is obtained by rewriting the terms in the original sum in the

form $x = qy + r$ and partitioning by congruence class to obtain

$$\begin{aligned} f(\alpha) &= \sum_{\substack{1\le x\le P\\ x=qy+r}} e\big((a/q+\beta)(qy+r)^k\big) \\ &= \sum_{r=1}^{q} e(ar^k/q) \sum_{y} e\big(\beta(qy+r)^k\big), \end{aligned}$$

where we have written $\beta = \alpha - a/q$. We then replace the sum over y by an integral over a continuous variable η, and make a change of variables to $\gamma = q\eta + r$ to obtain $q^{-1}\nu(\beta)$. We may do this at a cost of P^δ, which is negligible for a sufficiently small choice of δ. Consequently, $f(\alpha) \sim q^{-1}\nu(\beta)S(q,a)$ as claimed, and $f(\alpha)^s \sim q^{-s}\nu(\beta)^s S(q,a)^s$. Hence,

$$\int_{\mathfrak{M}(q,a)} f(\alpha)^s e(-n\alpha)\,d\alpha \sim q^{-s}S(q,a)^s e(-na/q) \int_{|\beta|<P^{\delta-k}} \nu(\beta)^s e(-n\beta)\,d\beta.$$

Combining the major arcs $\mathfrak{M}(q,a)$ for all $1 \le q \le Q$ and $1 \le a \le q$ with a coprime to q, we obtain a main term of the form

$$P^{s-k}\mathfrak{S}(P^\delta, n)J(P^\delta),$$

where

$$\mathfrak{S}(P^\delta, n) = \sum_{q\le P^\delta} \sum_{\substack{a=1\\(a,q)=1}}^{q} q^{-s}S(q,a)^s e(-an/q)$$

and

$$J(P^\delta) = \int_{|\beta|<P^\delta} \Big(\oint e(\beta\gamma^k)\,d\gamma \Big)^s e(-\beta)\,d\beta.$$

Examining these terms, we deduce that, whenever $s \ge k+1$, we have

$$\int_{\mathfrak{M}} f(\alpha)^s e(-\alpha n)\,d\alpha \sim \frac{\Gamma(1+1/k)^s}{\Gamma(s/k)} n^{s/k-1}\mathfrak{S}(n), \tag{1.1.3}$$

where

$$\Gamma(z) = \int_0^\infty t^{z-1}e^{-t}\,dt$$

is the classical Γ-function, and

$$\mathfrak{S}(n) = \sum_{q=1}^{\infty} \sum_{\substack{a=1 \\ (a,q)=1}}^{q} q^{-s} S(q,a)^s e(-an/q)$$

is known as the *singular series*, and is absolutely convergent for $s \geq 2k+1$ (see [21, Chapter 6]).

Turning to the minor arcs—namely, those points which lie on no major arc—we wish to show that their contribution to our integral is negligible. We present a number of auxiliary results which form part of the analysis of these arcs. Firstly, a well-known theorem of Dirichlet whose proof makes use of his pigeonhole principle.

Lemma 1.1.1 (Dirichlet's theorem, 1842)**.** *For $\alpha \in \mathbb{R}$ and $N \geq 1$, there exist $a, q \in \mathbb{Z}$ coprime with $1 \leq q \leq N$ such that*

$$|\alpha - a/q| \leq 1/(qN).$$

Secondly, we use a lemma originally due to Weyl in [62], and developed further by Hardy and Littlewood in [30].

Lemma 1.1.2 (Weyl's inequality, 1916)**.** *If $|\alpha - a/q| \leq 1/q^2$ for $a \in \mathbb{Z}$ and $q \in \mathbb{N}$ coprime, then*

$$f(\alpha) \ll P^{1+\epsilon}(q^{-1} + P^{-1} + qP^{-k})^{2^{1-k}}.$$

We adopt the convention throughout that statements involving ϵ hold for any suitably small choice of $\epsilon > 0$, and as such the exact value may change from line to line.

For $\alpha \in \mathfrak{m}$, Lemma 1.1.1 tells us that there exist $a, q \in \mathbb{N}$ with $q \leq Q^{-1}P^k$ and $|\alpha - a/q| \leq 1/(qQ^{-1}P^k) \leq QP^{-k}$. By the definition of $\mathfrak{m}$, we must therefore have $q > Q$, and since $1/(qQ^{-1}P^k) \leq 1/q^2$, we may also apply Lemma 1.1.2 to obtain

$$\begin{aligned} f(\alpha) &\ll P^{1+\epsilon}(Q^{-1} + P^{-1} + Q^{-1}P^kP^{-k})^{2^{1-k}} \\ &\ll P^{1+\epsilon}Q^{-2^{1-k}} \end{aligned} \tag{1.1.4}$$

for any $\alpha \in \mathfrak{m}$.

The following lemma, due to Hua in [36], allows us to obtain an improved upper bound for certain moments of our exponential sum.

Lemma 1.1.3 (Hua's lemma, 1938)**.** *For $1 \leq j \leq k$, we have*

$$\oint |f(\alpha)|^{2^j} \, d\alpha \ll P^{2^j - j + \epsilon}.$$

Then for $s > 2^k$, we have

$$\begin{aligned} \int_{\mathfrak{m}} f(\alpha)^s e(-\alpha n) \, d\alpha &\leq \int_{\mathfrak{m}} |f(\alpha)|^s \, d\alpha \\ &\ll \Big(\sup_{\alpha \in \mathfrak{m}} |f(\alpha)| \Big)^{s-2^k} \int_{\mathfrak{m}} |f(\alpha)|^{2^k} \, d\alpha \\ &\ll \big(P^{1+\epsilon} Q^{-2^{1-k}} \big)^{s-2^k} P^{2^k - k + \epsilon}, \end{aligned}$$

by (1.1.4) and Lemma 1.1.3, and so

$$\int_{\mathfrak{m}} f(\alpha)^s e(-\alpha n) \, d\alpha \ll \big(Q^{-2^{1-k}} \big)^{s-2^k} P^{s-k+\epsilon} = o(P^{s-k}).$$

We deduce that the equivalent of (1.1.3) holds with the integral on the left-hand side taken over the full unit interval. In order to conclude that solutions to our original equation (1.1.1) exist, it remains to show that $\mathfrak{S}(n) \gg 1$. It turns out that the singular series reflects the local solubility of the equation, in the following sense. Given a prime p, we say that an equation is *locally soluble at p* if it has a solution in the p-adic numbers $\mathbb{Q}_p$. It emerges that $\mathfrak{S}(n)$ is zero whenever the original equation is not locally soluble at some prime, which is unsurprising since this implies no integer solutions exist. Letting $\Gamma_0(k)$ denote the least value of s such that (1.1.1) has a non-singular solution in $\mathbb{Q}_p$ for all primes p, Hardy and Littlewood showed that $\mathfrak{S}(n) \gg 1$ whenever $s \geq \max\{\Gamma_0(k), 4\}$. They also proved that $\Gamma_0(k) = 4k$ for $k > 2$ a power of 2, and $\Gamma_0(k) \leq 2k$ otherwise, so certainly $\mathfrak{S}(n) \gg 1$ for $s \geq 4k$.

The above is a basic outline of the circle method as applied to Waring's problem, but it is possible to make further improvements via a related question concerning the number of solutions to the system of equations

$$x_1^j + \ldots + x_s^j = y_1^j + \ldots + y_s^j, \quad (1 \leq j \leq k), \tag{1.1.5}$$

with $1 \leq \boldsymbol{x}, \boldsymbol{y} \leq X$. Note that throughout, we use the vector notation $1 \leq \boldsymbol{x} \leq X$ to mean $1 \leq x_i \leq X$ for all i, as well as similar statements such as $\boldsymbol{x} \equiv \boldsymbol{y} \pmod{p}$ to mean $x_i \equiv y_i \pmod{p}$ for all i.

We denote the number of solutions to (1.1.5) with $1 \leq \boldsymbol{x}, \boldsymbol{y} \leq X$ by $J_{s,k}(X)$. An upper bound for $J_{s,k}(X)$ has become known as 'Vinogradov's Mean Value Theorem', with an optimal such bound being referred to as the 'Main Conjecture'. Vinogradov was interested in individual values of the exponential sums $\sum e(\alpha_1 x + \ldots + \alpha_k x^k)$, but studied them via their mean values

$$\oint \Big| \sum_{1 \leq x \leq X} e(\alpha_1 x + \ldots + \alpha_k x^k) \Big|^{2s} d\boldsymbol{\alpha};$$

by the orthogonality relation (1.1.2), the above integral is equal to $J_{s,k}(X)$.

In Waring's problem, we concern ourselves with an equation similar to that of highest degree in (1.1.5), and as such we may sum over all possible values for the expressions $\sum(x_i^j - y_i^j)$ for $1 \leq j \leq k-1$ to obtain a bound of the form

$$\int_{\mathfrak{m}} |f(\alpha)|^{2s} \, d\alpha \ll X^{k(k-1)/2} J_{s,k}(X),$$

and wish then to obtain the best possible bound for $J_{s,k}(X)$. Work of Vinogradov, Karatsuba and Stečkin (see, for example, [54]) leads to a result of the shape

$$J_{s,k}(X) \leq C(r,k) X^{2s - k(k+1)/2 + k^2(1-1/k)^r/2}$$

for $r \leq s/k$ a natural number.

Within the last few years, a significant body of work within number theory and harmonic analysis has provided an optimal upper bound for the number of solutions to the Vinogradov system (1.1.5), as follows.

Theorem 1.1.4. *For $k, s \in \mathbb{N}$, we have*

$$J_{s,k}(X) \ll X^{s+\epsilon} + X^{2s - k(k+1)/2}. \tag{1.1.6}$$

The case $k = 1$ is trivial, and when $k = 2$ we may use the quadratic identity

$$(a + b - c)^2 - (a^2 + b^2 - c^2) = 2(a - c)(b - c)$$

and the standard estimate for the divisor function, $d(n) \ll n^{\epsilon}$, to recover the

above bound. The case $k = 3$ is due to Wooley, who, in [69], obtained the first optimal upper bound for the number of solutions to (1.1.5) for any degree higher than two. His proof uses his method of *efficient congruencing*, first introduced in [67], and developed in a series of further papers—for an overview of the history of this method up until 2014 and its applications, see [68].

The central ideas of the efficient congruencing method revolve around the partition of our variables into congruence classes modulo some fixed prime p. We then use Hölder's inequality to restrict our variables to actually lie within a specified congruence class, and from here, we extract further congruence conditions on our variables which were previously 'hidden', and use these to lift our solutions, at a small cost, to a situation in which the variables are congruent modulo higher powers of p. Much of this is similar to the previous methodology, but the 'efficiency' of this method lies in the fact that, rather than using such a process to deduce statements of equality, we use it to generate continually stronger congruences modulo higher powers of p, which in turn generate stronger mean value estimates.

In parallel to the development of efficient congruencing, the technique of l^2-*decoupling* was coming to prominence in harmonic analysis, following the proof of the l^2-decoupling conjecture in [11] by Bourgain and Demeter. In [12], Bourgain, Demeter and Guth used this method to prove Theorem 1.1.4 in the general case $k \geq 4$, fully resolving the main conjecture of Vinogradov's mean value theorem. Wooley subsequently achieved the same result, with a number of further applications, in [71] via, specifically, the nested variant of his efficient congruencing method, which draws on ideas appearing previously in his work on discrete Fourier restriction estimates (see [70]). The techniques of decoupling and efficient congruencing are widely considered to be, respectively, real and p-adic versions of the same method, as discussed in [50].

1.2 Waring's problem with shifts

In the 1940s, Davenport and Heilbronn studied a generalisation of Waring's problem in which they considered diagonal inequalities of the form

$$\left|\lambda_1 x_1^k + \ldots + \lambda_s x_s^k\right| < 1, \tag{1.2.1}$$

where $\lambda_1 \dots \lambda_s \in \mathbb{R} \setminus \{0\}$ are not all of the same sign and not all in rational ratio. In [22], they proved that when $s \geq 2^k + 1$, there exist arbitrarily large P for which the number of integer solutions to (1.2.1) with $1 \leq x_1, \dots, x_s \leq P$ is at least cP^{s-k}, for some constant $c > 0$.

There are several differences between the method of Davenport and Heilbronn, and the original circle method of Hardy and Littlewood. Firstly, in order to count solutions to an inequality rather than an equation, we must integrate along the whole real line and use an appropriate kernel function to ensure convergence of the integral. Secondly, when evaluating the resulting integral, we observe that there is effectively only one major arc—one interval in which our function may be large—namely the arc around 0.

In [26] and [27], Freeman drew inspiration from work of Bentkus and Götze on quadratic forms (see [3]) to develop a version of the above method which delivers the same lower bound for the number of solutions, but for *all* large values of P. This has become known as Freeman's variant of the Davenport–Heilbronn method, and is now a crucial tool in the study of Diophantine inequalities. We will explore this further in Chapter 2, where we provide a more detailed introduction to both versions of the method.

In [15], Chow considered a different real analogue of Waring's problem, which has become known as *Waring's problem with shifts*. We fix natural numbers $s \geq k \geq 2$, 'shifts' $\theta_1, \dots, \theta_s \in (0, 1)$ with $\theta_1 \notin \mathbb{Q}$—this is the analogue of the requirement in the work of Davenport and Heilbronn that the coefficients not all be in rational ratio—and a small real number $\eta \in (0, 1]$. For a large, positive real number τ, we let $N(\tau) = N_{s,k,\boldsymbol{\theta},\eta}(\tau)$ be the number of solutions $(x_1, \dots, x_s) \in \mathbb{N}^s$ to the inequality

$$\left|(x_1 - \theta_1)^k + \dots + (x_s - \theta_s)^k - \tau\right| < \eta. \tag{1.2.2}$$

Chow used Freeman's variant of the Davenport–Heilbronn method to show that for $s \geq 2k^2 - 2k + 3$, we have the asymptotic formula

$$N(\tau) = 2\eta\Gamma(1 + 1/k)^s\Gamma(s/k)^{-1}\tau^{s/k-1} + o(\tau^{s/k-1}). \tag{1.2.3}$$

The recent proof of the main conjecture in Vinogradov's Mean Value Theorem, discussed in Section 1.1, allows the known range for the number of variables to be widened to $s \geq k^2 + k + 1$, and in Chapter 3, we further reduce

the minimum number of variables required. Let

$$s_0(k) = k^2 + (3k-1)/4.$$

Our main result in this area is the following, which first appeared in the author's paper [4]:

Theorem 1.2.1. *Let $k \geq 4$, and let $s \geq s_0(k)$. Then (1.2.3) holds.*

Note that our method also works for $k = 3$ and $s \geq 11$, but this simply recovers [14, Theorem 1.5], due to Chow.

The key tool leading to this improvement is a stronger bound on the contribution from the minor arcs, following the approach of Wooley in [66]. This reduction in the number of variables required on the minor arcs is sufficient to prove the result; in fact, Chow's argument requires only $k+1$ variables on the major arc, so we use his result in that case.

An interesting variant is to consider solutions of (1.2.2) lying in short intervals, which is to say, satisfying an additional condition of the form

$$X - Y < x_i \leq X + Y, \quad (1 \leq i \leq s), \tag{1.2.4}$$

with Y as small as possible, or equivalently

$$\left| x_i - (\tau/s)^{1/k} \right| < y(\tau), \quad (1 \leq i \leq s),$$

for some function $y(\tau)$, where we may think of $(\tau/s)^{1/k}$ as the 'average' value of our variables.

For the classical version of Waring's problem, Wright studied this question in [72], and proved that, if $\phi(n)$ is a function satisfying $\phi(n) \to 0$ as $n \to \infty$, then there exist arbitrarily large $n \in \mathbb{N}$ which cannot be represented in the form (1.1.1) subject to the additional condition $\left| x_i^k - n/s \right| < n^{1-1/2k}\phi(n)$ for $1 \leq i \leq s$. This result holds for all $s \in \mathbb{N}$, so this is an obstacle which cannot be overcome by increasing the number of variables.

However, widening the permissible region somewhat allows us to recover the existence of solutions. In [19], Daemen proves a lower bound on the number of representations of n under the condition

$$\left| x_i - (n/s)^{1/k} \right| < cn^{1/2k}, \quad (1 \leq i \leq s),$$

for a suitably large constant c, and in [18] he obtains an asymptotic formula under the condition

$$\left|x_i - (n/s)^{1/k}\right| < n^{1/2k+\epsilon}, \quad (1 \leq i \leq s).$$

In Section 3.6, which is based on the author's paper [5], we prove the following shifted analogue of Wright's result.

Theorem 1.2.2. *Let $s, k \geq 2$ be natural numbers. Fix $\boldsymbol{\theta} = (\theta_1, \dots, \theta_s) \in (0,1)^s$, and let $c, c' > 0$ be suitably small constants which may depend on s, k and $\boldsymbol{\theta}$. There exist arbitrarily large values of $\tau \in \mathbb{R}$ which cannot be approximated in the form (1.2.2), with $0 < \eta < c\tau^{1-2/k}$, subject to the additional condition that $\left|x_i - (\tau/s)^{1/k}\right| < c'\tau^{1/2k}$ for $1 \leq i \leq s$.*

In a sense this is surprising, since Waring's problem with shifts appears to concern real numbers, but it evidently retains some features of the classical problem—features which might have been supposed to be a consequence of the integer nature of the problem. The explanation for this rests on the spacing of the variables, or in other words the fact that the shifts θ_i are fixed, so that the kth powers $(x_i - \theta_i)^k$ which we consider are not entirely arbitrary.

1.3 Efficient congruencing with ellipsephic sets

In Chapters 4 and 5, we investigate variants of Vinogradov's mean value theorem in which the variables are drawn from thin subsets of the natural numbers satifying certain digital restrictions. Fix a subset $A \subset \mathbb{N}_0$, and a prime p. The sets we are interested in have the form

$$\mathcal{E} = \mathcal{E}_p^A = \{n \in \mathbb{N} \mid n = \textstyle\sum_i a_i p^i, \text{ with } a_i \in A \cap [0,p) \text{ for all } i\}.$$

In other words, we study the set of natural numbers whose digits in base p are drawn from A. We call such a set *ellipsephic*, a word which originated in the French mathematical literature (as *ellipséphique*), and appears to have been coined by Mauduit (see the discussion on page 12 of [17]), although such integers were studied prior to the introduction of the term—for example, in [25], Erdős, Mauduit and Sárközy studied the distribution of integers with missing digits in residue classes. Recent work of Maynard, in [46] and [47],

proves the existence of infinitely many primes with certain missing digits in a fixed base.

The cases where we permit no digits, or only the digit 0, are trivial, and that where we permit all digits reduces to the classical case, so we omit these from consideration. Furthermore, when we allow exactly one digit (which is not 0), the behaviour of $\mathcal{E}$ is different, with $\#\mathcal{E}(X) = \#\mathcal{E} \cap [1, X] \approx \log_p X$. As such, we restrict to the case where $2 \leq r = \#(A \cap [0, p)) \leq p - 1$, and note that we have

$$\#\mathcal{E}(X) \ll r^{\log_p X + 1} = rX^{\log_p r}.$$

Consequently, $\mathcal{E}$ is a thin subset of the integers, in the sense that

$$\lim_{X \to \infty} \frac{\#\mathcal{E}(X)}{X} = 0.$$

We observe that ellipsephic sets have a self-similar, fractal-like structure, with the digital restrictions seen here reminiscent of those in the classical middle-third Cantor set. They bear a resemblance to certain real fractal subsets constructed by Łaba and Pramanik in [39], for which those authors study maximal operators.

The bounds we obtain for the number of ellipsephic solutions to Vinogradov systems depend on the additive structure of the set A, and we now explain the specific property which interests us. A *Sidon set* is one in which all pairwise sums of elements from the set are distinct, and a *generalised Sidon set*, or *$B_h[g]$-set*, is one in which the number of representations of a natural number as the sum of h elements of the set is at most g, where representations are counted up to permutation of terms (so a Sidon set is a $B_2[1]$-set). For a survey of this area, see [48].

We generalise this concept further, as follows. For $t \geq 2$ an integer, and $\delta > 0$ a real number, we call a set $A \subset \mathbb{N}_0$ an $E_t(\delta)$-set if it has the property that

$$\#\{(a_1, \dots, a_t) \in A^t \mid a_1 + \dots + a_t = n\} \ll n^{\delta},$$

and an E_t^*-set if it is an $E_t(\delta)$-set for all positive δ.

It is desirable to have in mind examples of such sets. One such motivating case is the set of squares, which was shown by Landau in [41] to be an E_2^*-

set. In [31], Hardy and Littlewood made a conjecture known as Hypothesis K, which states that for all $k \geq 2$, the set of kth powers should be an E_k^*-set. However, in [44] Mahler proved that this hypothesis is false when $k = 3$, and it remains open for $k \geq 4$.

In [60], Vu showed that for any fixed $k \geq 2$, there exists a subset S_k of the set of kth powers and an integer t_k such that S_k is an $E_{t_k}^*$-set. This proves the existence of infinitely many sets of the form we are interested in, although the argument is probabilistic, so does not exhibit such sets directly.

We refer to a set $\mathcal{E}$ as a (p, t, δ)-ellipsephic set if $\mathcal{E} = \mathcal{E}_p^A$ and A is an $E_t(\delta)$-set, and as a $(p, t)^*$-ellipsephic set if A is an E_t^*-set. We now introduce some notation to allow us to state our main results. For a sequence $\mathfrak{a} = (\mathfrak{a}_n)_{n\in\mathcal{E}}$ of complex weights, we let

$$J_{s,k}(X; \mathfrak{a}) = \oint \Big| \sum_{x\in\mathcal{E}(X)} \mathfrak{a}_x e(\alpha_1 x + \ldots + \alpha_k x^k)\Big|^{2s} d\boldsymbol{\alpha},$$

which counts the solutions, in positive integers $x_i, y_i \in \mathcal{E}(X)$, to the system (1.1.5), where each solution is counted with weight $\mathfrak{a}_{\boldsymbol{x}}\overline{\mathfrak{a}_{\boldsymbol{y}}} = \mathfrak{a}_{x_1} \ldots \mathfrak{a}_{x_s}\overline{\mathfrak{a}_{y_1} \ldots \mathfrak{a}_{y_s}}$.

The main theorem of Chapter 4 provides the following upper bound for $J_{s,2}(X; \mathfrak{a})$.

Theorem 1.3.1. *For $t \geq 2$ an integer, $\delta > 0$ a real number, and p a suitably large prime, let $\mathcal{E}$ be a (p, t, δ)-ellipsephic set and let $Y = \#\mathcal{E}(X)$. Then for $s \geq 3t$, we have*

$$J_{s,2}(X; \mathfrak{a}) \ll Y^{s-3t} X^{3\delta+\epsilon} \bigg(\sum_{x\in\mathcal{E}(X)} |\mathfrak{a}_x|^2 \bigg)^s.$$

If $\mathcal{E}$ is a $(p, t)^*$-ellipsephic set, we therefore have

$$J_{s,2}(X; \mathfrak{a}) \ll Y^{s-3t} X^{\epsilon} \bigg(\sum_{x\in\mathcal{E}(X)} |\mathfrak{a}_x|^2 \bigg)^s.$$

The best upper bound which could previously be obtained for $J_{s,2}(X; \mathfrak{a})$ is a consequence of a result of Bourgain in [9]. Taking $\mathfrak{a}_x = 0$ for $x \notin \mathcal{E}$ in that

theorem yields, for $s \geq 3$,

$$J_{s,2}(X;\mathfrak{a}) \ll Y^{s-3}X^{\epsilon}\bigg(\sum_{x\in\mathcal{E}(X)}|\mathfrak{a}_x|^2\bigg)^s,$$

so we see that an improvement in the exponent of Y has been obtained by accounting for the specific structure of our ellipsephic sets, rather than just their density.

In the particular case of square digits, we have $t = 2$, so the conclusion of Theorem 1.3.1 at the critical case $s = 3t$ gives us

$$J_{6,2}(X;\mathfrak{a}) \ll X^{\epsilon}\bigg(\sum_{x\in\mathcal{E}(X)}|\mathfrak{a}_x|^2\bigg)^6,$$

as against the bound of

$$J_{6,2}(X;\mathfrak{a}) \ll Y^3X^{\epsilon}\bigg(\sum_{x\in\mathcal{E}(X)}|\mathfrak{a}_x|^2\bigg)^6$$

which follows from the work of Bourgain.

In Chapter 5, we develop this further to obtain the following bound for $J_{s,k}(X;\mathfrak{a})$ in the case of general k.

Theorem 1.3.2. *For natural numbers k and t, with $t \geq 2$, and p a suitably large prime, let $\mathcal{E}$ be a $(p,t)^*$-ellipsephic set and let $Y = \#\mathcal{E}(X)$. Then for $s \geq tk(k+1)/2$, we have*

$$J_{s,k}(X;\mathfrak{a}) \ll Y^{s-tk(k+1)/2}X^{\epsilon}\bigg(\sum_{x\in\mathcal{E}(X)}|\mathfrak{a}_x|^2\bigg)^s.$$

Again, in the case of square digits, we obtain

$$J_{k(k+1),k}(X;\mathfrak{a}) \ll X^{\epsilon}\bigg(\sum_{x\in\mathcal{E}(X)}|\mathfrak{a}_x|^2\bigg)^{k(k+1)}.$$

The proof of this theorem, as presented in Chapter 5, utilises the full power of Wooley's nested efficient congruencing method, with appropriate modifications to reflect the use of ellipsephic variables. Work similar to that in Chapter 4 provides the initial step in the inductive process.

1.4 Future work

1.4.1 Waring's problem with shifts

In [18], Daemen demonstrates an asymptotic formula for the number of solutions to the classical Waring's problem in which the variables are permitted to lie in short intervals slightly wider than those for which Wright proved that no solutions exist. Specifically, he requires the variables to satisfy a condition of the shape (1.2.4), and obtains an asymptotic formula for the number of such solutions whenever Y grows faster than $X^{1/2}$. Daemen's proof uses a method of binomial descent, in which one inductively bounds the number of solutions to the partial Vinogradov system

$$\sum_{i=1}^{s}(X+x_i)^k = \sum_{i=1}^{s}(X+y_i)^k,$$
$$\sum_{i=1}^{s}x_i^j = \sum_{i=1}^{s}y_i^j, \quad (1 \le j \le h),$$

for $0 \le h \le k-1$.

In future we plan to adapt Daemen's descent method to the case of shifted variables, and obtain an asymptotic formula for the number of solutions to (1.2.2) in short intervals.

1.4.2 Efficient congruencing with ellipsephic sets

One classical application of l^2-decoupling methods is to the d-dimensional paraboloid $(t_1, \dots, t_d, t_1^2 + \dots + t_d^2)$, considered by Bourgain and Demeter in [11]. As such, this is a natural problem to consider in our new ellipsephic setting. While the most straightforward case would be to draw all variables $t_1, \dots, t_d$ from the same ellipsephic set, the multidimensional nature of the problem allows the possibility of working with a d-dimensional ellipsephic set, where different restrictions apply to different variables.

Let $\mathcal{E}^{(1)}, \dots, \mathcal{E}^{(d)}$ be ellipsephic sets. We seek to estimate the number of

solutions to the system

$$x_{j,1} + \ldots + x_{j,s} = y_{j,1} + \ldots + y_{j,s}, \quad (1 \leq j \leq d),$$
$$x_{1,1}^2 + \ldots + x_{d,s}^2 = y_{1,1}^2 + \ldots + y_{d,s}^2,$$

with $x_{j,i}, y_{j,i} \in \mathcal{E}^{(j)}(X)$ for $1 \leq j \leq d$ and $1 \leq i \leq s$.

In the classical case, a bound for the number of solutions to a Vinogradov system such as (1.1.5) forms a key part of work on Waring's problem; consequently, this would be another natural extension of the work described in Chapters 4 and 5. In other words, we wish to represent all sufficiently large natural numbers n in the form (1.1.1) where $x_1, \ldots, x_s$ lie in some fixed ellipsephic set $\mathcal{E}$.

Given $k \geq 2$, we seek to find the smallest $s_0 = s_0(k)$ such that a representation of the above shape is possible, for sufficiently large n, whenever $s \geq s_0$, and the smallest $s_1 = s_1(k)$ for which we may obtain an asymptotic formula for the total number of such representations. There is an existing tradition of work on such problems for thin subsets of the integers, with the most commonly studied subset being the primes (see, for example, [37], or more recently [38]). In this case, our ellipsephic set $\mathcal{E}$ is much sparser than the primes, so there is the potential for results concerning thinner sets than any studied to date. A key challenge here will be the analysis of the main term: the smaller our set $\mathcal{E}$, the more difficult it becomes to control the error terms.

A more tractable form of this problem might be to seek solutions to

$$n = x_1^k + \ldots + x_s^k + y^k,$$

with $x_1, \ldots, x_s \in \mathcal{E}$ and $y \in \mathbb{N}_0$.

Finally, we mention a potential application of our work on ellipsephic sets to a different context, that of additive combinatorics. This is a relatively recent and growing area of mathematics, which links ideas and techniques from number theory, combinatorics, harmonic analysis and ergodic theory to study problems concerning the additive and multiplicative structure of the integers, and their interactions. For example, a key result in this area, due to

Roth in [52], states that any subset $\mathcal{A}$ of the natural numbers with

$$\limsup_{X\to\infty} \frac{\#(\mathcal{A}\cap[1,X])}{X} > 0 \tag{1.4.1}$$

contains a three-term arithmetic progression, or equivalently, a non-trivial solution to the equation $x+y=2z$.

In [13], Browning and Prendiville show that for $s \geq 5$, given a set of integers $c_1 \ldots, c_s$ with $c_1 + \ldots + c_s = 0$, any subset $\mathcal{A}$ of the natural numbers satisfying (1.4.1) contains a non-trivial solution to the equation

$$c_1 x_1^2 + \ldots + c_s x_s^2 = 0. \tag{1.4.2}$$

A corollary to this result is that the above equation is partition regular; that is, given any partition of the natural numbers into finitely-many colour classes, there exists a non-trivial monochromatic solution to (1.4.2). In [16], Chow generalises this by replacing the squares in (1.4.2) with kth powers, and shows that, for s at least as large as some $s_0 = s_0(k)$, any subset $\mathcal{P}$ of the primes with

$$\limsup_{X\to\infty} \frac{\#(\mathcal{P}\cap[1,X])}{X/\log X} > 0$$

contains a non-trivial solution to the resulting equation. Both of the aforementioned works use the transference technology of Green, introduced in [28], as a key ingredient.

There is the possibility of combining such problems from additive combinatorics with the ellipsephic ideas of Chapters 4 and 5, in order to prove Roth-type theorems in subsets of an ellipsephic set. The set $\mathcal{E}_2$ of integers whose digits in a fixed base are squares would be a key example. In this case, the appropriate equation to consider might be

$$x_1 + x_2 + x_3 + x_4 = 4x_5, \tag{1.4.3}$$

which is effectively the extension of the equation for a three-term arithmetic progression to the five-variable situation. The aim would therefore be to prove that any subset $\mathcal{A}$ of $\mathcal{E}_2$ with

$$\limsup_{X\to\infty} \frac{\#(\mathcal{A}\cap[1,X])}{\#\mathcal{E}_2(X)} > 0$$

contains a non-trivial solution to (1.4.3).

Chapter 2

The Davenport–Heilbronn Method

The aim of this chapter is to provide an introduction to the Davenport–Heilbronn method, and Freeman's variant thereof, for readers previously unfamiliar with them. We illustrate the method by giving sketch proofs of theorems of Davenport and Heilbronn, and of Freeman, demonstrating the key differences.

2.1 The classical method of Davenport and Heilbronn

In [22], Davenport and Heilbronn proved the following result.

Theorem 2.1.1. *Fix integers $k \geq 2$ and $s \geq 2^k + 1$. Let $\lambda_1, \ldots, \lambda_s$ be non-zero real numbers, not all of the same sign and not all in rational ratio. Then there exists a sequence $(P_n)_{n\in\mathbb{N}}$ with $P_n \to \infty$ as $n \to \infty$ with the property that for all $n \in \mathbb{N}$, the number of solutions $1 \leq x_1, \ldots, x_s \leq P_n$ to the inequality (1.2.1) is at least γP_n^{s-k} for some positive constant γ.*

In fact, Davenport and Heilbronn only prove Theorem 2.1.1 in detail for the case $k = 2$, but they remark that the same method will give the general case, a presentation of which may be found in Chapter 11 of [56].

Note that if $\lambda_1, \ldots, \lambda_s$ are all in rational ratio, one may find a real number λ_0 such that $\lambda_0\lambda_1, \ldots, \lambda_0\lambda_s$ are all integral, thereby reaching a Diophantine

equation; as such, we omit this case. Without loss of generality, we may assume that λ_1/λ_2 is irrational. We can also replace the right-hand side of the inequality (1.2.1) by any positive ϵ to reach the same conclusion, since we can apply the theorem with coefficients $\lambda_1/\epsilon, \dots, \lambda_s/\epsilon$.

We sketch the proof of this result in the case $s = 2^k + 1$. The key element is the introduction of the kernel function $K(\alpha) = \left(\frac{\sin(\pi\alpha)}{\pi\alpha}\right)^2$, which has the property (see [22, Lemma 4]) that for any real number t, we have

$$\int_{\mathbb{R}} e(t\alpha)K(\alpha)d\alpha = \max\{0, 1 - |t|\}. \tag{2.1.1}$$

For $\alpha \in \mathbb{R}$, let

$$f(\alpha) = f(\alpha, P) = \sum_{1 \le x \le P} e(\alpha x^k).$$

Then a (weighted) count of the number of solutions to (1.2.1) is given by

$$N(P) = \int_{\mathbb{R}} f(\lambda_1\alpha) \dots f(\lambda_s\alpha)K(\alpha)\, d\alpha.$$

We dissect the real line as follows. For large P, which will later be restricted to a certain sequence of values, and for fixed $\rho \in (0, 1/10)$, we define the major, minor and trivial arcs respectively as

$$\begin{aligned} \mathfrak{M} &= \{\alpha \in \mathbb{R} \mid |\alpha| \le P^{1-k}\}, \\ \mathfrak{m} &= \{\alpha \in \mathbb{R} \mid P^{1-k} < |\alpha| \le P^{\rho}\}, \end{aligned}$$

and

$$\mathfrak{t} = \{\alpha \in \mathbb{R} \mid P^{\rho} < |\alpha|\}.$$

On the major arc, we replace the sums $f(\alpha)$ by integrals of the form

$$\nu(\alpha) = \int_0^P e(\alpha\xi^k)\, d\xi.$$

By Euler's summation formula, we have that $f(\lambda_i\alpha) - \nu(\lambda_i\alpha) = O(1)$ whenever $|\alpha| \le P^{1/2-k}$. For the remainder of the major arc, where $P^{1/2-k} < |\alpha| \le P^{1-k}$, we use Dirichlet's theorem to find a rational approximation to α, and deduce

that $f(\alpha) = O(|\alpha|^{-1/2-\epsilon})$, and so

$$\begin{aligned}\int_{P^{1/2-k}<|\alpha|\le P^{1-k}} f(\lambda_1\alpha)\dots f(\lambda_s\alpha)K(\alpha)\,d\alpha &= O\bigg(\int_{P^{1/2-k}}^{\infty} \alpha^{-s(1/2+\epsilon)}\,d\alpha\bigg)\\ &= o(P^{s-k}).\end{aligned}$$

Combining these portions of the major arc, we conclude that

$$\int_{\mathfrak{M}} f(\lambda_1\alpha)\dots f(\lambda_s\alpha)K(\alpha)\,d\alpha = \int_{-\infty}^{\infty} \nu(\lambda_1\alpha)\dots\nu(\lambda_s\alpha)K(\alpha)\,d\alpha + o(P^{s-k}).$$

Due to the level of decay in $K(\alpha)$, we may now use Fubini's theorem to swap the order of integration on the right-hand side to obtain

$$\begin{aligned}&\int_{-\infty}^{\infty} \nu(\lambda_1\alpha)\dots\nu(\lambda_s\alpha)K(\alpha)\,d\alpha\\ &\quad= \int_0^P \dots \int_0^P \int_{-\infty}^{\infty} e\big(\alpha(\lambda_1\xi_1^k + \dots + \lambda_s\xi_s^k)\big)K(\alpha)\,d\alpha\,d\xi_1\dots\,d\xi_s\\ &\quad= \int_0^P \dots \int_0^P \max\{0, 1 - |\lambda_1\xi_1^k + \dots + \lambda_s\xi_s^k|\}\,d\xi_1\dots\,d\xi_s,\end{aligned}$$

by (2.1.1). We use a change of variables and rearrange to see that this is greater than a constant multiple of P^{s-k}, and therefore that

$$\int_{\mathfrak{M}} f(\lambda_1\alpha)\dots f(\lambda_s\alpha)K(\alpha)\,d\alpha \gg P^{s-k}, \tag{2.1.2}$$

as required.

It remains to show that the contributions from the minor and trivial arcs cannot possibly cancel out that from the major arc. By partitioning the integral into unit intervals, we may use Lemma 1.1.3 to conclude that

$$\int_{\mu}^{\infty} |f(\lambda_i\alpha)|^{2^k} K(\alpha)\,d\alpha = O\bigg(\frac{P^{2^k-k+\epsilon}}{\mu+1}\bigg) = O\bigg(\frac{P^{s-1-k+\epsilon}}{\mu+1}\bigg)$$

for $\mu \ge 0$ and $1 \le i \le s$. On the trivial arcs, this implies that

$$\int_{\mathfrak{t}} f(\lambda_1\alpha)\dots f(\lambda_s\alpha)K(\alpha)\,d\alpha = O(P^{s-k-\rho+\epsilon}) = o(P^{s-k}), \tag{2.1.3}$$

since we may always choose $\epsilon < \rho$.

To handle the minor arcs, we must now restrict P to lie in a certain sequence of integers, as mentioned above. Specifically, since λ_1/λ_2 is assumed to be irrational, Lemma 1.1.1 implies the existence of infinitely many coprime pairs $(a_0, q_0) \in \mathbb{Z} \times \mathbb{N}$ with

$$\left| \frac{\lambda_1}{\lambda_2} - \frac{a_0}{q_0} \right| \leq \frac{1}{q_0^2}.$$

We work henceforth with some choice of $P = q_0^2$. Using this assumption, we use Lemma 1.1.2 to deduce that

$$\min\{|f(\lambda_1\alpha)|, |f(\lambda_2\alpha)|\} = O(P^{1-\rho+\epsilon}) \tag{2.1.4}$$

whenever $\alpha \in \mathfrak{m}$. This step fails without the restriction on the value of P—in [56, Chapter 11], Vaughan claims that for certain λ_1 and λ_2,

$$\limsup_{P\to\infty} \left(\frac{1}{P} \sup_{\alpha\in\mathfrak{m}} \min\{|f(\lambda_1\alpha)|, |f(\lambda_2\alpha)|\} \right) > 0.$$

Specifically, one should take the ratio λ_1/λ_2 to be a Liouville number (a transcendental number which can be very closely approximated by rationals). Together with Lemma 1.1.3, we use (2.1.4) to conclude that

$$\int_{\mathfrak{m}} f(\lambda_1\alpha) \dots f(\lambda_s\alpha) K(\alpha)\, d\alpha = O(P^{s-k-\rho+\epsilon}) = o(P^{s-k}). \tag{2.1.5}$$

Theorem 2.1.1 follows by combining (2.1.2), (2.1.3) and (2.1.5).

2.2 Freeman's variant

In [26], Freeman adapted and developed the Davenport–Heilbronn method to deliver the above bound for all large values of P. He also reduced the required number of variables, and in [27] he proved an asymptotic formula for the number of solutions, but in this section we will focus on the lower bound in the case $s = 2^k + 1$, in order to draw clearer parallels with the method outlined in Section 2.1. The outline we provide here is based on the simplification of the method given by Wooley in [65].

The key improvement draws inspiration from [3], in which Bentkus and Götze study quadratic forms, to prove a bound of the following shape. Let $2 \leq S(P) \leq P$ be an increasing function tending to infinity with P, and let λ_1

and λ_2 be non-zero real numbers not in rational ratio. Then there exists an increasing function $T(P) \leq S(P)$ tending to infinity with P such that

$$\sup_{S(P)P^{-k} \leq |\alpha| \leq T(P)} |f(\lambda_1\alpha)f(\lambda_2\alpha)| \leq P^2 T(P)^{-2^{-k-1}}, \tag{2.2.1}$$

a bound which bears a noticeable similarity to (2.1.4). As in the original method described in Section 2.1, this is the point in the argument which requires λ_1/λ_2 to be irrational.

We now define the major, minor and trivial arcs in terms of the functions mentioned above, as

$$\begin{aligned} \mathfrak{M} &= \{\alpha \in \mathbb{R} \mid |\alpha| \leq S(P)P^{-k}\}, \\ \mathfrak{m} &= \{\alpha \in \mathbb{R} \mid S(P)P^{-k} < |\alpha| \leq T(P)\}, \end{aligned}$$

and

$$\mathfrak{t} = \{\alpha \in \mathbb{R} \mid T(P) < |\alpha|\}.$$

We also define a function $L(P) = \max\{1, \log T(P)\}$ which grows even more slowly than $T(P)$.

On the major arc, we follow a similar approach of replacing the sums $f(\lambda_i\alpha)$ with the integrals $v(\lambda_i\alpha)$, and using Fubini's theorem to deduce the required bound

$$\int_{\mathfrak{M}} f(\lambda_1\alpha) \dots f(\lambda_s\alpha)K(\alpha)\, d\alpha \gg P^{s-k}. \tag{2.2.2}$$

A more careful evaluation of the resulting integral yields the main term in an asymptotic formula for the number of solutions.

On the minor arcs, we subdivide further into regions reminiscent of the classical Hardy–Littlewood major and minor arcs. Let

$$\mathfrak{N} = \bigcup_{\substack{0 \leq a \leq q \leq S(P) \\ (a,q)=1}} \{\alpha \in [0,1) \mid |q\alpha - a| \leq S(P)P^{-k}\},$$

and $\mathfrak{n} = [0,1) \setminus \mathfrak{N}$. A standard result in this area (see [56, Chapter 4]) shows us that when $t > \max\{4, k+1\}$, we have

$$\int_{\mathfrak{N}} |f(\alpha)|^t \ll P^{t-k}, \tag{2.2.3}$$

and our choice of $s = 2^k + 1$ certainly satisfies this prerequisite. It is also the case, by [55, Theorem A], that

$$\int_{\mathfrak{n}} |f(\alpha)|^s \ll P^{s-k} T(P)^{-1}. \tag{2.2.4}$$

Consequently, for $n \in \mathbb{R}$ and $1 \leq i \leq s$, we have

$$\begin{aligned} \int_n^{n+1} |f(\lambda_i \alpha)|^s \, d\alpha &\ll \oint |f(\beta)|^s \, d\beta \\ &= \int_{\mathfrak{N}} |f(\beta)|^s \, d\beta + \int_{\mathfrak{n}} |f(\beta)|^s \, d\beta \ll P^{s-k}. \end{aligned} \tag{2.2.5}$$

We then partition the real line into

$$\mathfrak{P} = \{\alpha \in \mathbb{R} \mid \lambda_1 \alpha \pmod 1 \text{ lies in } \mathfrak{N}\},$$

and

$$\mathfrak{p} = \mathbb{R} \setminus \mathfrak{P} = \{\alpha \in \mathbb{R} \mid \lambda_1 \alpha \pmod 1 \text{ lies in } \mathfrak{n}\}.$$

For any unit interval $[n, n+1]$ contained in $\mathfrak{m}$, we use (2.2.4) to infer that

$$\int_{[n,n+1] \cap \mathfrak{p}} |f(\lambda_1 \alpha)|^s \ll \int_{\mathfrak{n}} |f(\lambda_1 \alpha)|^s \ll P^{s-k} T(P)^{-1},$$

and by Hölder's inequality, together with (2.2.5), we see that

$$\begin{aligned} \int_{[n,n+1] \cap \mathfrak{p}} |f(\lambda_1 \alpha) \dots f(\lambda_s \alpha)| \, d\alpha &\leq \prod_{i=1}^{s} \Big(\int_{[n,n+1] \cap \mathfrak{p}} |f(\lambda_i \alpha))|^s \, d\alpha \Big)^{1/s} \\ &\ll P^{s-k} T(P)^{-1/s} \ll P^{s-k} L(P)^{-2}. \end{aligned}$$

To handle $\mathfrak{P}$, we use (2.2.1) to see that for any unit interval $[n, n+1]$ contained in $\mathfrak{m}$, we have

$$\sup_{\alpha \in [n,n+1]} |f(\lambda_1 \alpha) f(\lambda_2 \alpha)| \leq P^2 T(P)^{-2^{-k-1}}.$$

Consequently, combining Hölder's inequality with (2.2.3) and (2.2.5), we obtain

$$\int_{[n,n+1]\cap\mathfrak{P}} |f(\lambda_1\alpha)\dots f(\lambda_s\alpha)|\, d\alpha \ll P^{s-k}T(P)^{-2^{-k-1}(s-2k)/(3s-2k)} \ll P^{s-k}L(P)^{-2}.$$

Summing the above bounds over all n with $[n, n+1] \subseteq \mathfrak{m}$, we deduce that

$$\int_{\mathfrak{m}} |f(\lambda_1\alpha)\dots f(\lambda_s\alpha)K(\alpha)|\, d\alpha \ll P^{s-k}L(P)^{-1} = o(P^{s-k}). \tag{2.2.6}$$

On the trivial arcs, we use (2.2.5) along with Hölder's inequality to obtain a bound

$$\int_n^{n+1} |f(\lambda_1\alpha)\dots f(\lambda_s\alpha)|\, d\alpha \ll P^{s-k}$$

for all real n, and then sum over the relevant values of $n \in \mathfrak{t}$ to conclude that

$$\int_{\mathfrak{t}} |f(\lambda_1\alpha)\dots f(\lambda_s\alpha)K(\alpha)|\, d\alpha \ll P^{s-k}L(P)^{-1} = o(P^{s-k}). \tag{2.2.7}$$

Again, the result follows from the combination of (2.2.2), (2.2.6) and (2.2.7).

Chapter 3

Waring's Problem with Shifts

The work in this chapter is based on the author's papers [4] and [5]. Sections 3.1–3.5 follow [4, Sections 1–5] respectively, and Section 3.6 follows [5].

3.1 Introduction

As discussed in Chapter 1, the classical form of Waring's problem asks whether every positive integer n can be represented as a sum of s kth powers of positive integers, where s does not depend on n. In Chapter 2, we considered a generalisation studied by Davenport and Heilbronn in which they sought solutions to the inequality (1.2.1) where the coefficients λ_i are non-zero, not all of the same sign, and not all in rational ratio.

As mentioned in Section 1.2, we now study a different analogue of Waring's problem, namely that of approximating real numbers by kth powers of shifted integers. More precisely, for a large, positive real number τ, we are interested in counting integer solutions to the inequality

$$\left|(x_1 - \theta_1)^k + \ldots + (x_s - \theta_s)^k - \tau\right| < \eta, \tag{3.1.1}$$

for fixed natural numbers $s \geq k \geq 2$, shifts $\theta_1, \ldots, \theta_s \in (0, 1)$ with θ_1 irrational, and $0 < \eta \leq 1$. Let $N(\tau) = N_{s,k,\boldsymbol{\theta},\eta}(\tau)$ be the number of solutions to (3.1.1) in positive integers $x_1, \ldots, x_s$. In this chapter, and the paper [4] from which it is taken, we reduce the minimum number of variables required to obtain an asymptotic formula for $N(\tau)$. To that end, let $s_0(k) = k^2 + (3k - 1)/4$. The main result of this chapter is the following theorem, presented in Chapter 1 as

Theorem 1.2.1, which we restate here for convenience.

Theorem 3.1.1. *Let $k \geq 4$, and let $s \geq s_0(k)$. Then*

$$N(\tau) = 2\eta\Gamma(1+1/k)^s\Gamma(s/k)^{-1}\tau^{s/k-1} + o(\tau^{s/k-1}). \tag{3.1.2}$$

Note that there is no explicit dependence on the shifts $\theta_1, \dots, \theta_s$ in the main term.

The best previously known bound for this problem is due to Chow, who showed in [15] that the asymptotic formula (3.1.2) holds for $s \geq 2k^2 - 2k + 3$. However, an examination of the arguments underlying Chow's work reveals that the recent proof in [12] of the Main Conjecture in Vinogradov's Mean Value Theorem, by Bourgain, Demeter and Guth, allows this constraint to be improved to $s \geq k^2 + k + 1$. Although our method also works for $k = 3$, it does not improve on the best known value of 11 variables, also due to Chow, in [14].

To prove our result, we use Freeman's variant of the Davenport–Heilbronn method, introduced in Section 2.2, which entails approximating the number of solutions to (3.1.1) by a certain integral over the real line, and using a dissection of the real line into major, minor and trivial arcs to evaluate this integral. However, in order to achieve our reduction in the number of variables required, we must also divide our arcs into points with or without good approximations by rationals with small denominators, commonly known as the major and minor arcs in the Hardy–Littlewood method.

The new estimate given in Section 3.3 extends the method of Wooley in [66] to the setting of Diophantine inequalities. We first obtain a bound for the contribution to a certain mean value from points without good rational approximations, making use of the aforementioned result of Bourgain, Demeter and Guth. In order to give a more precise statement of our result, we must introduce some notation. For real numbers P, θ and α, with P large and $\theta \in (0,1)$, we define

$$f_\theta(\alpha) = \sum_{1 \leq x \leq P} e(\alpha(x-\theta)^k).$$

We define $\mathfrak{v}$ to be the real analogue of the classical Hardy–Littlewood minor arcs: namely, with Q a real parameter satisfying $1 \leq Q \leq P$, we define $\mathfrak{v} = \mathfrak{v}_Q$

to be the set

$$\left\{\alpha \in \mathbb{R} : \big(a \in \mathbb{Z}, q \in \mathbb{N}, (a,q) = 1, |q\alpha - a| \leq QP^{-k}\big) \implies q > Q\right\}. \quad (3.1.3)$$

Finally, as in Chapter 2, we use the kernel function $K(\alpha) = \big(\frac{\sin(\pi\alpha)}{\pi\alpha}\big)^2$. We are now in a position to state the following result.

Theorem 3.1.2. *For natural numbers $s \geq 2$, $k \geq 2$, and for $\theta \in (0,1)$, we have*

$$\int_{\mathfrak{v}} |f_\theta(\alpha)|^{2s} K(\alpha)\, d\alpha \ll P^{\epsilon} Q^{-1} (P^{s+\frac{1}{2}k(k-1)} + P^{2s-k}). \quad (3.1.4)$$

This can be viewed as an analogue of the bound

$$\int_{\mathfrak{v}_{P/(2k)} \cap [0,1)} |f_0(\alpha)|^{2s}\, d\alpha \ll P^{\epsilon-1} (P^{s+\frac{1}{2}k(k-1)} + P^{2s-k}),$$

which is [66, Theorem 1.3], a result of Wooley.

We make use of a variant of Theorem 3.1.2 (see Theorem 3.3.1, and the subsequent conclusion in Corollary 3.3.6), which provides a key input to our application of Freeman's variant of the Davenport–Heilbronn method. The number of variables required to achieve this estimate is smaller than that required by Chow to bound the contribution from the same points, and this enables us to make our improvement as stated in Theorem 3.1.1. On the major arc, we use Chow's result to obtain the main term in the asymptotic formula, while on the remainder of the minor and trivial arcs not covered by Theorem 3.1.2, we show that the contribution is negligible. In order to do this, we make use of the measure of the set of points with good rational approximations, noting that these points constitute only a small fraction of any given unit interval.

We now present a brief outline of the structure of the remainder of this chapter. In Section 3.2, we introduce the preliminary notation required throughout the chapter. In Section 3.3, we present our new estimate for the contribution from the classical Hardy–Littlewood minor arcs, which ultimately allows us to improve on previously known lower bounds for the number of variables required for the asymptotic formula to hold. In Section 3.4, we show that negligible contributions are obtained from the remainder of the minor and trivial arcs not covered by Corollary 3.3.6. In Section 3.5 we present the result of

Chow on the major arc, giving the main term in the asymptotic formula for the number of solutions, thus completing the proof of Theorem 3.1.1. In Section 3.6 we consider the variant in which our variables are restricted to lie in short intervals.

3.2 Preliminary notation

We now introduce the conventions and pieces of standard notation which will be used in this chapter. Throughout, we assume that τ is sufficiently large in terms of $s, k, \boldsymbol{\theta} = (\theta_1, \dots, \theta_s)$ and η. Let $P = \tau^{1/k}$, and let $N^*(\tau)$ be the number of solutions to (3.1.1) with $1 \le \boldsymbol{x} \le P$. A solution which does not meet this condition can have at most one of the variables larger than $\tau^{1/k}$, and in this situation the remaining variables must each be at most some constant multiple of $\tau^{(k-1)/k^2}$. Thus, since we may assume that $s > k^2 - k + 1$, it follows that

$$N(\tau) - N^*(\tau) \ll \tau^{(s-1)(k-1)/k^2} = o(\tau^{s/k-1}).$$

It therefore suffices to prove that

$$N^*(\tau) = 2\eta\Gamma(1+1/k)^s\Gamma(s/k)^{-1}\tau^{s/k-1} + o(\tau^{s/k-1}).$$

We use the adapted kernel function

$$K(\alpha;\eta) = \eta\bigg(\frac{\sin(\pi\eta\alpha)}{\pi\eta\alpha}\bigg)^2,$$

which has the property (via a slight adaptation of [21, Lemma 20.1]) that for any real number t, we have

$$\int_{\mathbb{R}} e(t\alpha)K(\alpha;\eta)\,d\alpha = \max\{0, 1 - |t/\eta|\}. \tag{3.2.1}$$

Consequently, letting

$$f_{\boldsymbol{\theta}}(\alpha) = f_{\theta_1}(\alpha)\cdots f_{\theta_s}(\alpha),$$

we observe that the integral

$$\int_{\mathbb{R}} f_{\boldsymbol{\theta}}(\alpha)e(-\tau\alpha)K(\alpha;\eta)\,d\alpha \tag{3.2.2}$$

provides a weighted count of the number of solutions to (3.1.1). To be precise, a tuple $(x_1, \dots, x_s)$ contributes 1 whenever the left-hand side of (3.1.1) is equal to zero, and $1 - \zeta/\eta$ whenever the left-hand side of (3.1.1) is equal to ζ, for some $\zeta \in (0, \eta)$.

The following lemma, which is similar to the result presented in (2.2.1), demonstrates the existence of a certain positive function which provides a bound on the values of the exponential sums we are interested in.

Lemma 3.2.1. *Let $k \geq 2$ be an integer, and let $\xi, \theta_1, \theta_2 \in (0,1)$ with θ_1 irrational. Then there exists a positive real-valued function $T(P)$ with $T(P) \to \infty$ as $P \to \infty$, such that*

$$\sup_{P^{\xi-k} \leq |\alpha| \leq T(P)} |f_{\theta_1}(\alpha) f_{\theta_2}(\alpha)| \ll P^2 T(P)^{-1}. \tag{3.2.3}$$

Proof. This is a special case of [15, Lemma 2.2]. □

We divide up the real line into major, minor and trivial arcs, as is usual in the Davenport–Heilbronn method. We fix a real number $\xi \in (0,1)$, and apply Lemma 3.2.1 to obtain the function $T(P)$. We then define

$$\begin{aligned} \mathfrak{M} &= \{\alpha \in \mathbb{R} : |\alpha| \leq P^{\xi-k}\}, \\ \mathfrak{m} &= \{\alpha \in \mathbb{R} : P^{\xi-k} < |\alpha| \leq T(P)\}, \end{aligned}$$

and

$$\mathfrak{t} = \{\alpha \in \mathbb{R} : |\alpha| > T(P)\}.$$

We can therefore evaluate the integral (3.2.2) using the dissection

$$\mathbb{R} = \mathfrak{M} \cup \mathfrak{m} \cup \mathfrak{t}. \tag{3.2.4}$$

In order to successfully evaluate the contribution from the central major arc (as in [15, Section 3]), we must use a different kernel function related to $K(\alpha; \eta)$ to reduce the length of the interval which provides a non-negligible contribution. We define

$$L(P) = \min\{\log T(P), \log P\}, \quad \delta = \eta L(P)^{-1}, \tag{3.2.5}$$

and the upper and lower kernel functions

$$K_{\pm}(\alpha) = \frac{\sin(\pi\alpha\delta)\sin(\pi\alpha(2\eta \pm \delta))}{\pi^2\alpha^2\delta}.$$

These kernel functions are the same as those obtained in [27, Lemma 1] (applied with $a = \eta - \delta$ and $b = \eta$ for $K_-(\alpha)$, and with $a = \eta$ and $b = \eta + \delta$ for $K_+(\alpha)$, along with $h = 1$ in both cases). Letting $U_c(t)$ denote the indicator function of the interval $(-c, c)$, the conclusion of that lemma gives us the bounds

$$U_{\eta-\delta}(t) \leq \int_{\mathbb{R}} e(\alpha t)K_-(\alpha)d\alpha \leq U_\eta(t),$$

$$U_\eta(t) \leq \int_{\mathbb{R}} e(\alpha t)K_+(\alpha)d\alpha \leq U_{\eta+\delta}(t),$$

and

$$K_{\pm}(\alpha) \ll \min\{1, \alpha^{-2}L(P)\}. \tag{3.2.6}$$

Letting

$$R_{\pm}(P) = \int_{\mathbb{R}} f_{\boldsymbol{\theta}}(\alpha)e(-\tau\alpha)K_{\pm}(\alpha)\, d\alpha,$$

we therefore have

$$R_-(P) \leq N^*(\tau) \leq R_+(P).$$

Consequently, it suffices to prove that

$$R_{\pm}(P) = 2\eta\Gamma(1 + 1/k)^s\Gamma(s/k)^{-1}P^{s-k} + o(P^{s-k}).$$

In the approximations which follow, we need to use estimates for the above kernel functions, and as such it is helpful to note the following decomposition (see [49, Section 2]). We write

$$|K_{\pm}(\alpha)|^2 = K_1(\alpha)K_2^{\pm}(\alpha), \tag{3.2.7}$$

where

$$K_1(\alpha) = \left(\frac{\sin(\pi\alpha\delta)}{\pi\alpha\delta}\right)^2 = \delta^{-1}K(\alpha;\delta) \tag{3.2.8}$$

and

$$K_2^{\pm}(\alpha) = \left(\frac{\sin(\pi\alpha(2\eta \pm \delta))}{\pi\alpha}\right)^2 = (2\eta \pm \delta)K(\alpha; 2\eta \pm \delta) \tag{3.2.9}$$

are both non-negative.

Using (3.2.1), we also note that

$$\int_{\mathbb{R}} K_1(\alpha)e(\alpha t)\,d\alpha = \begin{cases} \delta^{-1}(1-\delta^{-1}|t|), & \text{if } |t|<\delta, \\ 0, & \text{otherwise,} \end{cases} \tag{3.2.10}$$

and

$$\int_{\mathbb{R}} K_2^{\pm}(\alpha)e(\alpha t)\,d\alpha = \begin{cases} 2\eta \pm \delta - |t|\,, & \text{if } |t|<2\eta\pm\delta, \\ 0, & \text{otherwise.} \end{cases} \tag{3.2.11}$$

3.3 An auxiliary estimate

In this section, we achieve a bound on the contribution from the traditional Hardy–Littlewood minor arcs, namely those points which are not close to a rational number with small denominator. In doing so, we improve on Chow's result for the number of variables required for the asymptotic formula (3.1.2) to hold. We follow closely the method of Wooley in [66, Section 2]. Thus, we firstly obtain an estimate for a related mean value, defined below in (3.3.4), in the case where we have a single shift $\theta = \theta_1 = \ldots = \theta_s$. We then use this result, along with Hölder's inequality, to bound the integral we are interested in, and to generalise to the case in which the shifts need not be the same.

It is convenient to introduce some further notation for use in this section. We define the exponential sums

$$g(\boldsymbol{\alpha}) = g_k(\boldsymbol{\alpha},\theta;P) = \sum_{1\le x\le P} e(\alpha_1 x + \ldots + \alpha_{k-1}x^{k-1} + \alpha_k(x-\theta)^k),$$

and

$$G(\boldsymbol{\beta},\mu) = G_k(\boldsymbol{\beta},\mu,\theta;P) = \sum_{1\le x\le P} e(\beta_1 x + \ldots + \beta_{k-2}x^{k-2} + \mu(x-\theta)^k),$$

as well as the polynomials

$$\sigma_{s,j}(\mathbf{x}) = \sum_{i=1}^{s}(x_i^j - x_{s+i}^j), \quad (1\le j\le k-1),$$

and

$$\sigma_{s,k}(\mathbf{x}) = \sum_{i=1}^{s}((x_i - \theta)^k - (x_{s+i} - \theta)^k).$$

We also define the integral

$$I_{s,k}^{\pm}(P,\theta) = \int_{\mathbb{R}} |f_\theta(\alpha)|^{2s} |K_\pm(\alpha)| \, d\alpha. \tag{3.3.1}$$

Let $J_{s,k}(P,\theta)$ be the number of solutions of the system

$$\begin{cases} \sigma_{s,j}(\mathbf{x}) = 0, \quad (1 \le j \le k-1), \\ |\sigma_{s,k}(\mathbf{x})| < \eta, \end{cases} \tag{3.3.2}$$

with $1 \le \boldsymbol{x} \le P$. Using binomial expansions, and the fact that $\eta \le 1$, we see that this system is equivalent to the system of Diophantine equations

$$\begin{cases} \sigma_{s,j}(\mathbf{x}) = 0, \quad (1 \le j \le k-1), \\ \sum_{i=1}^{s}(x_i^k - x_{s+i}^k) = 0, \end{cases} \tag{3.3.3}$$

which is precisely the Vinogradov system (1.1.5) discussed in Chapter 1. It is therefore the case that $J_{s,k}(P,\theta) = J_{s,k}(P)$, the number of solutions to (3.3.3) with $1 \le \boldsymbol{x} \le P$.

For $1 \le Q \le P$, we define $\mathfrak{v} = \mathfrak{v}_Q$ as in (3.1.3). In later applications we will consider $Q = (2k)^{-1}P^{1/4}$. We are interested in an estimate for the minor arc portion (in the Hardy–Littlewood sense) of the integral $I_{s,k}^{\pm}(P,\theta)$ defined in (3.3.1). For $B \subset \mathbb{R}$ measurable, we write

$$I^{\pm}(B) = I_{s,k}^{\pm}(B,P,\theta) = \int_B |f_\theta(\alpha)|^{2s} |K_\pm(\alpha)| \, d\alpha. \tag{3.3.4}$$

This allows us to state the key result of this section.

Theorem 3.3.1. *For natural numbers $s \ge 2$, $k \ge 2$, and for $\theta \in (0,1)$, we have*

$$I^{\pm}(\mathfrak{v}) \ll P^{\epsilon}Q^{-1}(P^{s+\frac{1}{2}k(k-1)} + P^{2s-k}).$$

Proof. We would like to rewrite the integral of interest in terms of the function $G(\boldsymbol{\beta},\mu)$, in order to separate out the x^{k-1} term and estimate it using the rational approximation properties of points in $\mathfrak{v}$.

For $\mathbf{h} \in \mathbb{Z}^{k-2}$, let

$$\delta(\mathbf{x}, \mathbf{h}) = \prod_{j=1}^{k-2} \Big(\oint e\big(\beta_j(\sigma_{s,j}(\mathbf{x}) - h_j)\big)\, d\beta_j \Big),$$

which, by orthogonality, is equal to 1 if $\sigma_{s,j}(\mathbf{x}) = h_j$ for all $1 \le j \le k-2$, and zero otherwise.

For any fixed $\mathbf{x} \in [1, P]^{2s}$, there is precisely one choice of $\mathbf{h} \in \mathbb{Z}^{k-2}$ which satisfies the above condition, and by the definition of $\sigma_{s,j}$ we have $|\sigma_{s,j}(\mathbf{x})| \le sP^j$ for $1 \le j \le k-2$. Hence

$$\sum_{|h_1| \le sP} \dots \sum_{|h_{k-2}| \le sP^{k-2}} \delta(\mathbf{x}, \mathbf{h}) = 1.$$

We can therefore rewrite the minor arc integral in the form

$$\begin{aligned} I^{\pm}(\mathfrak{v}) &= \int_{\mathfrak{v}} \sum_{1 \le \mathbf{x} \le P} e(\mu \sigma_{s,k}(\mathbf{x}))\, |K_{\pm}(\mu)|\, d\mu \\ &= \sum_{\boldsymbol{h}} \sum_{1 \le \mathbf{x} \le P} \delta(\mathbf{x}, \mathbf{h}) \int_{\mathfrak{v}} e(\mu \sigma_{s,k}(\mathbf{x}))\, |K_{\pm}(\mu)|\, d\mu, \end{aligned}$$

where the first summation is over $(k-2)$-tuples $\boldsymbol{h}$ satisfying $|h_i| \le sP^i$ for $1 \le i \le k-2$. From the definition of $G(\boldsymbol{\beta}, \mu)$, we obtain

$$\begin{aligned} I^{\pm}(\mathfrak{v}) &= \sum_{\boldsymbol{h}} \sum_{1 \le \mathbf{x} \le P} \prod_{j=1}^{k-2} \Big(\oint e\big(\beta_j(\sigma_{s,j}(\mathbf{x}) - h_j)\big)\, d\beta_j \Big) \int_{\mathfrak{v}} e(\mu \sigma_{s,k}(\mathbf{x}))\, |K_{\pm}(\mu)|\, d\mu \\ &= \sum_{\boldsymbol{h}} \int_{\mathfrak{v}} \oint |G(\boldsymbol{\beta}, \mu)|^{2s}\, e(-\boldsymbol{\beta} \cdot \boldsymbol{h})\, |K_{\pm}(\mu)|\, d\boldsymbol{\beta}\, d\mu. \end{aligned}$$

Hence, using the triangle inequality, and defining

$$\mathcal{I} = \int_{\mathfrak{v}} \oint |G(\boldsymbol{\beta}, \mu)|^{2s}\, |K_{\pm}(\mu)|\, d\boldsymbol{\beta}\, d\mu, \tag{3.3.5}$$

we see that

$$\begin{aligned} I^{\pm}(\mathfrak{v}) &\le \sum_{\boldsymbol{h}} \int_{\mathfrak{v}} \oint |G(\boldsymbol{\beta}, \mu)|^{2s}\, |K_{\pm}(\mu)|\, d\boldsymbol{\beta}\, d\mu \\ &\ll P^{\frac{1}{2}(k-1)(k-2)} \mathcal{I}. \end{aligned} \tag{3.3.6}$$

Similarly, writing

$$g(\boldsymbol{\alpha},\mu)=\sum_{1\le x\le P} e(\alpha_1 x+\ldots+\alpha_{k-1}x^{k-1}+\mu(x-\theta)^k),$$

we have

$$\mathcal{I}=\sum_{|h|\le sP^{k-1}}\int_{\mathfrak{v}}\oint |g(\boldsymbol{\alpha},\mu)|^{2s}\, e(-\alpha_{k-1}h)\,|K_{\pm}(\mu)|\,d\boldsymbol{\alpha}\,d\mu. \tag{3.3.7}$$

Let $\psi(z;\boldsymbol{\alpha})=\alpha_1 z+\ldots+\alpha_{k-1}z^{k-1}+\alpha_k(z-\theta)^k$, so that, with a change of variables, we have

$$g(\boldsymbol{\alpha})=\sum_{1\le x\le P} e(\psi(x;\boldsymbol{\alpha}))=\sum_{1+y\le x\le P+y} e(\psi(x-y;\boldsymbol{\alpha})). \tag{3.3.8}$$

Define

$$\mathcal{L}(\gamma)=\sum_{1\le z\le P} e(-\gamma z),$$

and

$$\mathfrak{g}_y(\boldsymbol{\alpha};\gamma)=\sum_{1\le x\le 2P} e\big(\psi(x-y;\boldsymbol{\alpha})+\gamma(x-y)\big),$$

so that

$$\overline{\mathfrak{g}}_y(\boldsymbol{\alpha};\gamma)=\mathfrak{g}_y(-\boldsymbol{\alpha};-\gamma).$$

Then, for $1\le y\le P$, we observe from (3.3.8) that

$$\begin{aligned}\oint \mathfrak{g}_y(\boldsymbol{\alpha};\gamma)\mathcal{L}(\gamma)\,d\gamma &= \oint \sum_{1\le x\le 2P}\sum_{1\le z\le P} e\big(\psi(x-y;\boldsymbol{\alpha})+\gamma(x-y-z)\big)\,d\gamma\\ &= \sum_{1\le x\le 2P}\sum_{\substack{1\le z\le P\\ z=x-y}} e\big(\psi(x-y;\boldsymbol{\alpha})\big)\\ &= g(\boldsymbol{\alpha}).\end{aligned}$$

Substituting this relation into (3.3.7), we find that

$$\mathcal{I}=\sum_{|h|\le sP^{k-1}}\int_{\mathfrak{v}}\oint \Big|\oint \mathfrak{g}_y(\boldsymbol{\alpha},\mu;\gamma)\mathcal{L}(\gamma)\,d\gamma\Big|^{2s} e(-\alpha_{k-1}h)\,|K_{\pm}(\mu)|\,d\boldsymbol{\alpha}\,d\mu.$$

Writing

$$\mathcal{G}_y(\boldsymbol{\alpha},\mu;\boldsymbol{\gamma})=\prod_{i=1}^{s}\mathfrak{g}_y(\boldsymbol{\alpha},\mu;\gamma_i)\overline{\mathfrak{g}}_y(\boldsymbol{\alpha},\mu;\gamma_{s+i}),$$

and

$$\tilde{\mathcal{L}}(\boldsymbol{\gamma})=\prod_{i=1}^{s}\mathcal{L}(\gamma_i)\mathcal{L}(-\gamma_{s+i}),$$

we see that

$$\mathcal{I}=\sum_{|h|\le sP^{k-1}}\int_{\mathfrak{v}}\oint\cdots\oint\mathcal{G}_y(\boldsymbol{\alpha},\mu;\boldsymbol{\gamma})\tilde{\mathcal{L}}(\boldsymbol{\gamma})e(-\alpha_{k-1}h)\,|K_{\pm}(\mu)|\;d\boldsymbol{\gamma}\,d\boldsymbol{\alpha}\,d\mu.$$

If we let

$$I_h(\boldsymbol{\gamma},y)=\int_{\mathfrak{v}}\oint\cdots\oint\mathcal{G}_y(\boldsymbol{\alpha},\mu;\boldsymbol{\gamma})e(-\alpha_{k-1}h)\,|K_{\pm}(\mu)|\;d\boldsymbol{\alpha}\,d\mu, \tag{3.3.9}$$

then we can write

$$\mathcal{I}=\sum_{|h|\le sP^{k-1}}\oint I_h(\boldsymbol{\gamma},y)\tilde{\mathcal{L}}(\boldsymbol{\gamma})\,d\boldsymbol{\gamma}. \tag{3.3.10}$$

Evaluating the inner integral of $I_h(\boldsymbol{\gamma},y)$ using orthogonality, we see that

$$\oint\cdots\oint\mathcal{G}_y(\boldsymbol{\alpha},\mu;\boldsymbol{\gamma})e(-\alpha_{k-1}h)\,|K_{\pm}(\mu)|\;d\boldsymbol{\alpha}=|K_{\pm}(\mu)|\sum_{1\le\mathbf{x}\le 2P}\Delta_{\boldsymbol{x}}, \tag{3.3.11}$$

where

$$\Delta_{\boldsymbol{x}}=\Delta_{\boldsymbol{x}}(\mu,\boldsymbol{\gamma},h,y)=e\Big(\mu\sigma_{s,k}(\mathbf{x}-y)+\sum_{i=1}^{s}\big(\gamma_i(x_i-y)-\gamma_{s+i}(x_{s+i}-y)\big)\Big)$$

whenever

$$\begin{cases}\sigma_{s,j}(\mathbf{x}-y)=0, & (1\le j\le k-2),\\ \sigma_{s,k-1}(\mathbf{x}-y)=h,\end{cases} \tag{3.3.12}$$

and otherwise $\Delta_{\boldsymbol{x}}(\mu,\boldsymbol{\gamma},h,y)=0$.

Using binomial expansions, we see that whenever the above conditions (3.3.12) hold, we also have the relations

$$\sigma_{s,j}(\mathbf{x})=0=\sigma_{s,j}(\mathbf{x}-\theta),$$

for $1 \le j \le k-2$, and

$$\sigma_{s,k-1}(\mathbf{x}) = h = \sigma_{s,k-1}(\mathbf{x} - \theta),$$

and consequently

$$\sigma_{s,k}(\mathbf{x} - y) = \sum_{i=1}^{s} \big((x_i - \theta - y)^k - (x_{s+i} - \theta - y)^k \big) = \sigma_{s,k}(\mathbf{x}) - khy.$$

We therefore see from (3.3.11) that

$$\oint \mathcal{G}_y(\boldsymbol{\alpha}, \mu; \boldsymbol{\gamma}) e(-\alpha_{k-1}h) \, |K_{\pm}(\mu)| \, d\boldsymbol{\alpha} \\ = \oint |K_{\pm}(\mu)| \, \mathcal{G}_0(\boldsymbol{\alpha}, \mu; \boldsymbol{\gamma}) e(-\mu khy - \alpha_{k-1}h) \, \omega_{y,\boldsymbol{\gamma}} \, d\boldsymbol{\alpha},$$

where $\omega_{y,\boldsymbol{\gamma}} = e(-y\sigma_{s,1}(\boldsymbol{\gamma}))$. Using (3.3.9), we have

$$\sum_{|h| \le sP^{k-1}} I_h(\boldsymbol{\gamma}, y) \\ = \int_{\mathfrak{v}} \oint |K_{\pm}(\mu)| \, \mathcal{G}_0(\boldsymbol{\alpha}, \mu; \boldsymbol{\gamma}) \sum_{|h| \le sP^{k-1}} e(-\mu khy - \alpha_{k-1}h) \, \omega_{y,\boldsymbol{\gamma}} \, d\boldsymbol{\alpha} \, d\mu \\ \ll \int_{\mathfrak{v}} \oint |K_{\pm}(\mu)| \, |\mathcal{G}_0(\boldsymbol{\alpha}, \mu; \boldsymbol{\gamma})| \min\{P^{k-1}, \|\mu ky + \alpha_{k-1}\|^{-1}\} \, d\boldsymbol{\alpha} \, d\mu$$

by a standard geometric series estimate for exponential sums (see, for example, [21, Chapter 3]).

Averaging over all permitted values of y, and writing

$$\Psi(\mu, \alpha_{k-1}) = P^{-1} \sum_{1 \le y \le P} \min\{P^{k-1}, \|\mu ky + \alpha_{k-1}\|^{-1}\}, \tag{3.3.13}$$

we see that

$$P^{-1} \sum_{1 \le y \le P} \sum_{|h| \le sP^{k-1}} I_h(\boldsymbol{\gamma}, y) \\ \ll \int_{\mathfrak{v}} \oint |K_{\pm}(\mu)| \, |\mathcal{G}_0(\boldsymbol{\alpha}, \mu; \boldsymbol{\gamma})| \, \Psi(\mu, \alpha_{k-1}) \, d\boldsymbol{\alpha} \, d\mu. \tag{3.3.14}$$

Now we find a rational approximation for μ. By Lemma 1.1.1, there exist

$b \in \mathbb{Z}$ and $r \in \mathbb{N}$ with $(b, r) = 1$ such that $r \leq P^k Q^{-1}$ and $|r\mu - b| \leq QP^{-k} \leq r^{-1}$. We make use of a lemma of Baker, namely [1, Lemma 3.2], which tells us that, under these approximation conditions, we have

$$\begin{aligned}\sum_{1 \leq y \leq P} \min\{P^{k-1}, \|\mu y + \alpha_{k-1}\|^{-1}\} &\ll \big(P^{k-1} + r \log(2r)\big)(Pr^{-1} + 1) \\ &\ll P^k(r^{-1} + P^{-1} + rP^{-k}) \log(2r).\end{aligned}$$

A minor modification to the proof of this result allows us to incorporate the additional factor of k in (3.3.13) and conclude that

$$\Psi(\mu, \alpha_{k-1}) \ll P^{k-1}(r^{-1} + P^{-1} + rP^{-k}) \log(2r).$$

By the definition of $\mathfrak{v}$, we have $r > Q$, and therefore

$$\sup_{\mu \in \mathfrak{v}} \Psi(\mu, \alpha_{k-1}) \ll Q^{-1} P^{k-1} \log P.$$

Substituting this into (3.3.14) and using Hölder's inequality, we see that

$$\begin{aligned}P^{-1} &\sum_{1 \leq y \leq P} \sum_{|h| \leq sP^{k-1}} I_h(\boldsymbol{\gamma}, y) \\ &\ll Q^{-1} P^{k-1} (\log P) \int_{\mathfrak{v}} \oint |K_{\pm}(\mu)| \left| \prod_{i=1}^{s} \mathfrak{g}_0(\boldsymbol{\alpha}, \mu; \gamma_i) \overline{\mathfrak{g}}_0(\boldsymbol{\alpha}, \mu; \gamma_{s+i}) \right| d\boldsymbol{\alpha}\, d\mu \\ &\ll Q^{-1} P^{k-1} (\log P) \prod_{i=1}^{2s} \left(\int_{\mathfrak{v}} \oint |K_{\pm}(\mu)| \, |\mathfrak{g}_0(\boldsymbol{\alpha}, \mu; \gamma_i)|^{2s} \, d\boldsymbol{\alpha}\, d\mu \right)^{1/2s} \\ &\ll Q^{-1} P^{k-1} (\log P) \sup_{\gamma \in [0,1)} \int_{\mathbb{R}} \oint |K_{\pm}(\mu)| \, |\mathfrak{g}_0(\boldsymbol{\alpha}, \mu; \gamma)|^{2s} \, d\boldsymbol{\alpha}\, d\mu \\ &\ll Q^{-1} P^{k-1} (\log P) \int_{\mathbb{R}} \oint |K_{\pm}(\mu)| \, |g_k(\boldsymbol{\alpha}, \mu, \theta; 2P)|^{2s} \, d\boldsymbol{\alpha}\, d\mu. \qquad (3.3.15)\end{aligned}$$

For a general function $H \colon \mathbb{R} \to \mathbb{R}$, we write

$$\Upsilon(H) = \int_{\mathbb{R}} |H(\mu)| \, |g_k(\boldsymbol{\alpha}, \mu, \theta; 2P)|^{2s} \, d\mu.$$

Using the Cauchy–Schwarz inequality, and the decomposition (3.2.7), we ob-

tain

$$\oint \Upsilon(K_{\pm})\, d\boldsymbol{\alpha} \le \Big(\oint \Upsilon(K_1)\, d\boldsymbol{\alpha} \Big)^{1/2} \Big(\oint \Upsilon(K_2^{\pm})\, d\boldsymbol{\alpha} \Big)^{1/2}.$$

From (3.2.10), we deduce that $\Upsilon(K_1)$ contributes

$$\delta^{-1}(1 - \delta^{-1}\,|\sigma_{s,k}(\boldsymbol{x})|)\, e(\alpha_1\sigma_{s,1}(\boldsymbol{x}) + \ldots + \alpha_{k-1}\sigma_{s,k-1}(\boldsymbol{x}))$$

whenever $|\sigma_{s,k}(\boldsymbol{x})| < \delta$. Recalling that $\delta = \eta L(P)^{-1} \le \eta$ for sufficiently large P, and using the equivalence of systems (3.3.2) and (3.3.3), this implies that

$$\oint \Upsilon(K_1)\, d\boldsymbol{\alpha} \le \delta^{-1} J_{s,k}(2P) \ll L(P) J_{s,k}(2P).$$

Similarly, using (3.2.11), we have

$$\oint \Upsilon(K_2^{\pm})\, d\boldsymbol{\alpha} \ll J_{s,k}(2P).$$

We remark that we also have $\oint \Upsilon(K)\, d\boldsymbol{\alpha} \ll J_{s,k}(2P)$, which allows us to establish the simplified claim (3.1.4) given in Section 3.1.

Substituting the above estimates into (3.3.15) and using (3.2.5), we see that

$$P^{-1} \sum_{1 \le y \le P} \sum_{|h| \le sP^{k-1}} I_h(\boldsymbol{\gamma}, y) \ll Q^{-1} P^{k-1} (\log P)^{3/2}\, J_{s,k}(2P).$$

Returning to (3.3.10), and noting that $\mathcal{I}$ as originally defined in (3.3.5) does not depend on y, we see that

$$\mathcal{I} = P^{-1} \sum_{1 \le y \le P} \mathcal{I} \ll Q^{-1} P^{k-1} (\log P)^{3/2}\, J_{s,k}(2P) \oint \Big| \tilde{\mathcal{L}}(\boldsymbol{\gamma}) \Big|\, d\boldsymbol{\gamma}. \tag{3.3.16}$$

By the definition of $\mathcal{L}(\gamma)$, we have

$$\oint |\mathcal{L}(\gamma)|\, d\gamma \le \oint \min\{P, \|\gamma\|^{-1}\}\, d\gamma \ll \log P,$$

and therefore

$$\oint \Big| \tilde{\mathcal{L}}(\boldsymbol{\gamma}) \Big|\, d\boldsymbol{\gamma} = \oint \left| \prod_{i=1}^{s} \mathcal{L}(\gamma_i) \mathcal{L}(-\gamma_{s+i}) \right| d\boldsymbol{\gamma} \ll (\log P)^{2s}.$$

Substituting this into (3.3.16), we see that

$$\mathcal{I} \ll Q^{-1}P^{k-1}(\log P)^{2s+3/2}\, J_{s,k}(2P),$$

and hence, from (3.3.6), that

$$\begin{aligned} I^{\pm}(\mathfrak{v}) &\ll Q^{-1}P^{\frac{1}{2}k(k-1)}(\log P)^{2s+3/2}\, J_{s,k}(2P) \\ &\ll Q^{-1}P^{\frac{1}{2}k(k-1)+\epsilon}\, J_{s,k}(2P). \end{aligned}$$

Using (1.1.6), we conclude that

$$I^{\pm}(\mathfrak{v}) \ll P^{\epsilon}Q^{-1}(P^{s+\frac{1}{2}k(k-1)} + P^{2s-k}),$$

as required. □

In particular, we have

$$\int_{\mathfrak{v}} |f_\theta(\mu)|^{2s}\, |K_{\pm}(\mu)|\, d\mu \ll Q^{-1}P^{s+\frac{1}{2}k(k-1)+\epsilon}$$

whenever $s \leq \frac{1}{2}k(k+1)$, and

$$\int_{\mathfrak{v}} |f_\theta(\mu)|^{2s}\, |K_{\pm}(\mu)|\, d\mu \ll Q^{-1}P^{2s-k+\epsilon}$$

whenever $s \geq \frac{1}{2}k(k+1)$.

We now wish to use the above result to bound the minor arc contribution for our shifted Waring's problem. From this point onwards, we fix $Q = (2k)^{-1}P^{1/4}$. We use the Cauchy–Schwarz inequality and a trivial estimate in order to limit the number of variables needed to achieve the required bound, which ultimately allows us to prove Theorem 3.1.1 (in Section 3.5). We then go on to provide a conjectural further improvement (for $k = 10$ and $k \geq 12$) based on an adaptation of a theorem of Bourgain (arising from the results in [12]).

Corollary 3.3.2. *Let $k \geq 2$ be a natural number, and let $s_0(k) = k^2 + (3k-1)/4$. Then for any real number $s \geq s_0(k)$, we have*

$$\int_{\mathfrak{v}} |f_\theta(\alpha)|^{s}\, |K_{\pm}(\alpha)|\, d\alpha = o(P^{s-k}). \tag{3.3.17}$$

Proof. Fix $k \geq 2$, and let $s_0 = s_0(k)$. We first prove (3.3.17) in the case $s = s_0$. We wish to apply Hölder's inequality, so we set

$$a = \frac{s_0 - 2}{(k+2)(k-1)}, \quad b = \frac{k^2 + k - s_0}{(k+2)(k-1)}.$$

Note that by the definition of s_0, and since $k > 5/3$, we have

$$b = \frac{k+1}{4k^2 + 4k - 8} < \frac{k+1}{4k^2 + k - 3} = \frac{1}{4k-3}, \tag{3.3.18}$$

which will be crucial in the conclusion of the proof. We have $a + b = 1$, and $ak(k+1) + 2b = s_0$, so, using the notation introduced in (3.3.4), and suppressing the dependence on k, P and θ, we can apply Hölder's inequality to see that

$$\int_{\mathfrak{v}} |f_\theta(\alpha)|^{s_0} |K_\pm(\alpha)| \, d\alpha \ll \big(I^\pm_{k(k+1)/2}(\mathfrak{v})\big)^a \big(I^\pm_1(\mathfrak{v})\big)^b.$$

We evaluate the first term using Theorem 3.3.1 to get

$$I^\pm_{k(k+1)/2}(\mathfrak{v}) \ll Q^{-1} P^{k(k+1)-k+\epsilon}.$$

For the second term, we use the decomposition (3.2.7), along with the Cauchy–Schwarz inequality, to obtain

$$I^\pm_1(\mathfrak{v}) \ll \bigg(\int_{\mathfrak{v}} |f_\theta(\alpha)|^2 K_1(\alpha) \, d\alpha \bigg)^{1/2} \bigg(\int_{\mathfrak{v}} |f_\theta(\alpha)|^2 K_2^\pm(\alpha) \, d\alpha \bigg)^{1/2}.$$

Since the number of solutions to the inequality $\big|(x-\theta)^k - (y-\theta)^k\big| < \delta$ with $1 \leq x, y \leq P$ is $O(P)$, we use (3.2.10) and (3.2.5) to see that

$$\int_{\mathfrak{v}} |f_\theta(\alpha)|^2 K_1(\alpha) \, d\alpha \leq \int_{\mathbb{R}} |f_\theta(\alpha)|^2 K_1(\alpha) \, d\alpha \ll L(P) P.$$

Similarly, using (3.2.11), we have

$$\int_{\mathfrak{v}} |f_\theta(\alpha)|^2 K_2^\pm(\alpha) \, d\alpha \leq \int_{\mathbb{R}} |f_\theta(\alpha)|^2 K_2^\pm(\alpha) \, d\alpha \ll P.$$

We therefore see that

$$I_1^{\pm}(\mathfrak{v}) \ll (\log P)^{1/2} P \ll P^{1+\epsilon}.$$

Hence, with some rearrangement, and using the definitions of a, b and Q,

$$\begin{aligned}\int_{\mathfrak{v}} |f_\theta(\alpha)|^{s_0} |K_\pm(\alpha)| \, d\alpha &\ll P^{a(k(k+1)-k-1/4+\epsilon)+b(1+\epsilon)} \\ &= P^{s_0-k+\epsilon-\iota},\end{aligned}$$

where

$$\iota = a/4 - b(k-1) = 1/4 - b(k-3/4) > \epsilon$$

for small enough ϵ, by (3.3.18).

For $s > s_0$, we then use the trivial estimate to obtain

$$\begin{aligned}\int_{\mathfrak{v}} |f_\theta(\alpha)|^{s} |K_\pm(\alpha)| \, d\alpha &\ll P^{s-s_0} \int_{\mathfrak{v}} |f_\theta(\alpha)|^{s_0} |K_\pm(\alpha)| \, d\alpha \\ &\ll P^{s-s_0} P^{s_0-k+\epsilon-\iota} = o(P^{s-k}).\end{aligned}$$

□

We now present a more sophisticated version of the above argument, which follows a similar structure. For $j < k$ a natural number, we define

$$\begin{aligned}s_1(k,j) &= \left\lceil k(k+1) - \frac{k(k+1)-j(j+1)}{4(k-j)+1} \right\rceil + 1 \\ &= k^2 + k + 1 - \left\lfloor \frac{k(k+1)-j(j+1)}{4(k-j)+1} \right\rfloor.\end{aligned}$$

We require an improved version of Lemma 1.1.3. Since [12, Theorem 4.1] applies equally to the case of exponential sums of suitably separated points, such as the set $\{x - \theta : x \in \mathbb{N}\}$, as it does to the integer case, it would seem that the following 'shifted' analogue of [10, Theorem 10] should hold. However, the details of such a result do not yet appear in the literature.

Hypothesis 3.3.3 ('Shifted Hua's Lemma'). *For $j \leq k$ a natural number, and for any fixed, positive ζ, we have*

$$\int_{\mathbb{R}} |f_\theta(\alpha)|^{j(j+1)} K(\alpha;\zeta) \, d\alpha \ll P^{j^2+\epsilon}. \tag{3.3.19}$$

Note that the implicit constant in (3.3.19) may depend on ζ.

Corollary 3.3.4. *Assuming the shifted Hua's lemma, for any natural number $s \geq s_1(k,j)$ we have*

$$\int_{\mathfrak{v}} |f_\theta(\alpha)|^s \, |K_\pm(\alpha)| \, d\alpha = o(P^{s-k}). \tag{3.3.20}$$

Proof. Fix j and k, and let $s_1 = s_1(k,j)$. We first prove (3.3.20) in the case $s = s_1$. Let

$$a = \frac{s_1 - j(j+1)}{k(k+1) - j(j+1)}, \; b = \frac{k(k+1) - s_1}{k(k+1) - j(j+1)}.$$

Note that by the definition of s_1 we have

$$b = \frac{k(k+1) - \left\lceil k(k+1) - \frac{k(k+1)-j(j+1)}{4(k-j)+1} \right\rceil - 1}{k(k+1) - j(j+1)} < \frac{1}{4(k-j)+1}. \tag{3.3.21}$$

We have $a + b = 1$, and $ak(k+1) + bj(j+1) = s_1$, so, as in Corollary 3.3.2, we can apply Hölder's inequality to see that

$$\int_{\mathfrak{v}} |f_\theta(\alpha)|^{s_1} \, |K_\pm(\alpha)| \, d\alpha \ll \left(I^\pm_{k(k+1)/2}(\mathfrak{v})\right)^a \left(I^\pm_{j(j+1)/2}(\mathfrak{v})\right)^b.$$

We evaluate the first term using Theorem 3.3.1 to get

$$I^\pm_{k(k+1)/2}(\mathfrak{v}) \ll Q^{-1} P^{k(k+1)-k+\epsilon}.$$

For the second term, as in Corollary 3.3.2, we obtain

$$\begin{aligned} I^\pm_{j(j+1)/2}(\mathfrak{v}) &\ll \left(\int_{\mathfrak{v}} |f_\theta(\alpha)|^{j(j+1)} \, K_1(\alpha) \, d\alpha \right)^{1/2} \left(\int_{\mathfrak{v}} |f_\theta(\alpha)|^{j(j+1)} \, K_2^\pm(\alpha) \, d\alpha \right)^{1/2} \\ &\ll \left(\int_{\mathbb{R}} |f_\theta(\alpha)|^{j(j+1)} \, K_1(\alpha) \, d\alpha \right)^{1/2} \left(\int_{\mathbb{R}} |f_\theta(\alpha)|^{j(j+1)} \, K_2^\pm(\alpha) \, d\alpha \right)^{1/2}. \end{aligned}$$

Combining (3.2.8) and (3.2.9) with the shifted Hua's lemma, we see that

$$\int_{\mathbb{R}} |f_\theta(\alpha)|^{j(j+1)} \, K_1(\alpha) \, d\alpha \ll L(P) P^{j^2+\epsilon} \ll P^{j^2+\epsilon},$$

and

$$\int_{\mathbb{R}} |f_\theta(\alpha)|^{j(j+1)} K_2^{\pm}(\alpha)\, d\alpha \ll P^{j^2+\epsilon},$$

and therefore that

$$I^{\pm}_{j(j+1)/2}(\mathfrak{v}) \ll P^{j^2+\epsilon}.$$

Hence, with some rearrangement, and using the definitions of a, b and Q,

$$\begin{aligned}\int_{\mathfrak{v}} |f_\theta(\alpha)|^{s_1} |K_\pm(\alpha)|\, d\alpha &\ll P^{(k(k+1)-k+\epsilon-1/4)a+(j(j+1)-j+\epsilon)b} \\ &= P^{s_1-k+\epsilon-\iota}\end{aligned}$$

where $\iota = 1/4 - (k-j+1/4)b > \epsilon$ for small enough ϵ, by (3.3.21).

For $s > s_1$, we then use the trivial estimate to obtain

$$\begin{aligned}\int_{\mathfrak{v}} |f_\theta(\alpha)|^{s} |K_\pm(\alpha)|\, d\alpha &\ll P^{s-s_1} \int_{\mathfrak{v}} |f_\theta(\alpha)|^{s_1} |K_\pm(\alpha)|\, d\alpha \\ &\ll P^{s-s_1} P^{s_1-k+\epsilon-\iota} \\ &= o(P^{s-k}).\end{aligned}$$

□

Optimisation shows that, for a given k, the minimal value of $s_1(k,j)$ occurs when

$$j = j_0(k) = \left[k + \frac{1}{4} - \sqrt{\frac{1}{2}k + \frac{5}{16}}\right]$$

where $[x]$ denotes the nearest integer to x. Note that for all $k \geq 2$, we have $j_0(k) < k$. Letting $s_1(k) = s_1(k, j_0(k))$, we therefore conclude the following corollary, noting that $s_1(k) = k^2 + k/2 + O(k^{1/2})$, and that $s_1(k) < s_0(k)$ for $k = 10$ and $k \geq 12$.

Corollary 3.3.5. *Assuming the shifted Hua's lemma, for any natural number $s \geq s_1(k)$, we have*

$$\int_{\mathfrak{v}} |f_\theta(\alpha)|^{s} |K_\pm(\alpha)|\, d\alpha = o(P^{s-k}).$$

Finally, we generalise the above results to the case of mixed shifts $\theta_1, \dots, \theta_s$.

Corollary 3.3.6. *Suppose that $\theta_1, \dots, \theta_s \in (0,1)$, and write $\boldsymbol{\theta} = (\theta_1, \dots, \theta_s)$.*

Then for any natural number $s \geq s_0(k)$, we have

$$\int_{\mathfrak{v}} |f_{\boldsymbol{\theta}}(\alpha)|\,|K_{\pm}(\alpha)|\,d\alpha = o(P^{s-k}).$$

Assuming the shifted Hua's lemma, the same result holds whenever $s \geq s_1(k)$.

Proof. By Hölder's inequality, and using Corollary 3.3.2 or Corollary 3.3.5, as appropriate, we have

$$\begin{aligned}\int_{\mathfrak{v}} |f_{\boldsymbol{\theta}}(\alpha)|\,|K_{\pm}(\alpha)|\,d\alpha &\ll \prod_{i=1}^{s}\left(\int_{\mathfrak{v}} |f_{\theta_i}(\alpha)|^s\,|K_{\pm}(\alpha)|\,d\alpha\right)^{1/s}\\ &= o(P^{s-k}).\end{aligned}$$

□

3.4 The minor and trivial arcs

On the minor and trivial arcs, we first demonstrate the estimate

$$\int_{\mathfrak{m}\cup\mathfrak{t}} \left|f_{\theta_1}(\alpha)f_{\theta_2}(\alpha)f_{\theta_3}(\alpha)^{s-2}K_{\pm}(\alpha)\right|\,d\alpha = o(P^{s-k}),$$

for shifts $\theta_1, \theta_2, \theta_3 \in (0,1)$ with θ_1 irrational, and later use Hölder's inequality to obtain the general case. We subdivide our arcs into those points with good rational approximations and those without, in a manner reminiscent of the classical Hardy–Littlewood method, by defining

$$\begin{aligned}\mathfrak{N}_{a,q} &= \{\alpha \in \mathfrak{m}\cup\mathfrak{t} : |q\alpha - a| \leq QP^{-k}\},\\ \mathfrak{N} &= \bigcup_{\substack{1\leq q\leq Q,\\ (a,q)=1}} \mathfrak{N}_{a,q}, \quad \text{and} \quad \mathfrak{n} = (\mathfrak{m}\cup\mathfrak{t})\setminus\mathfrak{N}.\end{aligned}$$

Those points in $\mathfrak{n}$ are handled using Corollary 3.3.6, since $\mathfrak{n} = \mathfrak{v}\cap(\mathfrak{m}\cup\mathfrak{t})$, while those in $\mathfrak{N}$ are subdivided yet again on the basis of the size of the exponential sum $f_{\theta_3}(\alpha)$. For some real number t (to be chosen later) satisfying $2k(k-1)t < 1$, let

$$\mathfrak{B} = \mathfrak{B}(t) = \{\alpha \in \mathfrak{N} : |f_{\theta_3}(\alpha)| \geq P^{1-t}\},$$

and

$$\overline{\mathfrak{B}} = \overline{\mathfrak{B}}(t) = \mathfrak{N} \setminus \mathfrak{B},$$

so that

$$\mathfrak{m} \cup \mathfrak{t} = \mathfrak{B} \cup \overline{\mathfrak{B}} \cup \mathfrak{n}.$$

Let $\mathfrak{B}_v$, $\overline{\mathfrak{B}}_v$ denote the intersection of $\mathfrak{B}$, $\overline{\mathfrak{B}}$ respectively with the unit interval $[v, 1+v)$. Note that for any $v \in \mathbb{R}$, we have

$$\operatorname{mes}(\mathfrak{B}_v \cup \overline{\mathfrak{B}}_v) \leq \sum_{q=1}^{Q} \sum_{a=1}^{q} 2QP^{-k}/q \ll Q^2 P^{-k}. \tag{3.4.1}$$

We use this to bound the contribution to the overall integral from $\overline{\mathfrak{B}}$.

Lemma 3.4.1. *Let t be such that $2k(k-1)t < 1$. For $u > 1/(2t)$, and for any $v \in \mathbb{R}$, we have*

$$\int_{\overline{\mathfrak{B}}_v} |f_{\theta_3}(\alpha)|^u \, d\alpha = o(P^{u-k}).$$

Proof. Note that, by assumption, we have $ut > 1/2$. Therefore, using (3.4.1), and recalling that $Q = (2k)^{-1}P^{1/4}$,

$$\begin{aligned}\int_{\overline{\mathfrak{B}}_v} |f_{\theta_3}(\alpha)|^u \, d\alpha &\ll (P^{1-t})^u \operatorname{mes}(\overline{\mathfrak{B}}_v) \ll P^{u-ut} Q^2 P^{-k} \\ &\ll P^{u-k+1/2-ut} = o(P^{u-k}).\end{aligned}$$

□

We therefore have the following estimate for those $\overline{\mathfrak{B}}_v$ contained in the minor arcs.

Lemma 3.4.2. *For $s \geq k^2 + 2$, and for any v with $\overline{\mathfrak{B}}_v \subset \mathfrak{m}$, there exists $\iota > 0$ such that*

$$\int_{\overline{\mathfrak{B}}_v} \left| f_{\theta_1}(\alpha) f_{\theta_2}(\alpha) f_{\theta_3}(\alpha)^{s-2} \right| d\alpha \ll P^{s-k-\iota} T(P)^{-1}.$$

Proof. By choosing t so that $2k(k-1)t$ is as close as we like to 1, note that we can always find a u such that $1/(2t) < u < k^2 \leq s-2$. Applying Lemma

3.4.1 and (3.2.3), we see that

$$\begin{aligned}\int_{\overline{\mathfrak{B}}_v} \left|f_{\theta_1}(\alpha)f_{\theta_2}(\alpha)f_{\theta_3}(\alpha)^{s-2}\right| d\alpha \\ &\ll \sup_{\alpha\in\overline{\mathfrak{B}}_v} \left|f_{\theta_1}(\alpha)f_{\theta_2}(\alpha)f_{\theta_3}(\alpha)^{s-2-u}\right| \int_{\overline{\mathfrak{B}}_v} |f_{\theta_3}(\alpha)|^u \, d\alpha \\ &\ll P^2\, T(P)^{-1}(P^{1-t})^{s-2-u}P^{u-k} \\ &\ll P^{s-k-(s-2-u)t}\, T(P)^{-1} = P^{s-k-\iota}\, T(P)^{-1},\end{aligned}$$

where $\iota = (s-2-u)t > 0$. □

Consequently, we can add in the contribution from the trivial arcs to show that we have the required estimate on $\overline{\mathfrak{B}}$.

Lemma 3.4.3. *We have*

$$\int_{\overline{\mathfrak{B}}} \left|f_{\theta_1}(\alpha)f_{\theta_2}(\alpha)f_{\theta_3}(\alpha)^{s-2}K_{\pm}(\alpha)\right| d\alpha = o(P^{s-k}).$$

Proof. Combining Lemma 3.4.2 with (3.2.6) and (3.4.1), we find that

$$\begin{aligned}&\int_{\overline{\mathfrak{B}}} \left|f_{\theta_1}(\alpha)f_{\theta_2}(\alpha)f_{\theta_3}(\alpha)^{s-2}K_{\pm}(\alpha)\right| d\alpha \\ &\quad\ll \sum_{v=0}^{\infty} \int_{\overline{\mathfrak{B}}_{v+P^{\xi-k}}} \left|f_{\theta_1}(\alpha)f_{\theta_2}(\alpha)f_{\theta_3}(\alpha)^{s-2}K_{\pm}(\alpha)\right| d\alpha \\ &\quad\ll (T(P)-P^{\xi-k})P^{s-k-\iota}\, T(P)^{-1} + \sum_{v=-1}^{\infty} \frac{L(P)}{(v+T(P))^2} P^2(P^{1-t})^{s-2}P^{1/2-k}.\end{aligned}$$

Since $1/2 < t(s-2)$ by our choice of t, we conclude that

$$\begin{aligned}\int_{\overline{\mathfrak{B}}} \left|f_{\theta_1}(\alpha)f_{\theta_2}(\alpha)f_{\theta_3}(\alpha)^{s-2}K_{\pm}(\alpha)\right| d\alpha &\ll P^{s-k-\iota} + \frac{L(P)}{T(P)-1} P^{s-k+1/2-t(s-2)} \\ &= o(P^{s-k}).\end{aligned}$$

□

On $\mathfrak{B}$, we use the result [2, Theorem 4] of Baker, which improves on the earlier result [67, Theorem 1.6] of Wooley, in the light of [12]. Firstly, we define $\alpha_0, \dots, \alpha_k$ by

$$\alpha(x-\theta_3)^k = \sum_{i=0}^{k} \alpha_i x^i, \tag{3.4.2}$$

and note that $\alpha_k = \alpha$.

Theorem 3.4.4. *Let $k \geq 3$ be an integer, and let t be a positive real number with $2k(k-1)t < 1$. Let ζ be a sufficiently small positive real number. Suppose that P is sufficiently large, and that $|f_{\theta_3}(\alpha)| \geq P^{1-t}$. Then there exist integers $q, a_1, \ldots, a_k$ such that $1 \leq q \leq P^{1-\zeta}$ and $|q\alpha_j - a_j| \leq P^{1-j-\zeta}$ for $1 \leq j \leq k$. These integers also satisfy $(q, a_1, \ldots, a_k) = 1$ and $(q, a_2, \ldots, a_k) \leq 2k^2$.*

Proof. This is the case $A = P^{1-t}$ of [2, Theorem 4]. The direct conclusion is that for small λ, we have $1 \leq q \leq P^{\lambda+kt}$ and $|q\alpha_j - a_j| \leq P^{-j+\lambda+kt}$ for $1 \leq j \leq k$; by choosing λ such that $2(k-1)(\lambda+t) < 1$ and $0 < \zeta < 1-\lambda-kt$, we reach the conclusion given above. It is also possible to extract from the proof of this result in [2] that the greatest common divisor $d = (q, a_2, \ldots, a_k)$ satisfies $d \leq 2k^2$. Restricting to $(q, a_1, \ldots, a_k) = 1$ can only reduce the values of q and d, so nothing is lost by doing so. □

We introduce some more notation. For integers $q, a_1, \ldots, a_k$, write

$$S(q, \mathbf{a}) = \sum_{x=1}^{q} e\Big(\frac{a_k x^k + \ldots + a_1 x}{q}\Big),$$

and for real numbers $\beta_1, \ldots, \beta_k$, write

$$I(\boldsymbol{\beta}) = \int_0^P e(\beta_k y^k + \ldots + \beta_1 y)\, dy.$$

We will use Theorem 3.4.4 in conjunction with the following lemma.

Lemma 3.4.5. *Let $k \geq 2$. Let $f(x) = \alpha_k x^k + \ldots + \alpha_1 x$, and suppose that there are integers $q, a_1, \ldots, a_k$ such that*

$$|q\alpha_j - a_j| \leq (2k^2)^{-1} P^{1-j}, \quad (1 \leq j \leq k).$$

Writing

$$d = (q, a_2, \ldots, a_k),$$

and

$$\beta_j = \alpha_j - \frac{a_j}{q}, \quad (1 \leq j \leq k),$$

we have

$$\sum_{x=1}^{P} e(f(x)) = q^{-1} S(q, \mathbf{a}) I(\boldsymbol{\beta}) + O(q^{1-1/k+\epsilon} d^{1/k}).$$

Proof. This is [1, Lemma 4.4]. □

By the definition of $\mathfrak{B}$, we have met the conditions of Theorem 3.4.4 for $\alpha \in \mathfrak{B}$. Fixing a sufficiently small $\zeta > 0$, and a choice of λ with $2(k-1)(\lambda+t) < 1$ and $0 < \zeta < 1 - \lambda - kt$, we may let $q(\alpha), a_1(\alpha), \dots, a_k(\alpha)$ be integers meeting the conditions given in the conclusion of that theorem, namely that $1 \leq q(\alpha) \leq P^{1-\zeta}$ and $|q(\alpha)\alpha_j - a_j(\alpha)| \leq P^{1-j-\zeta}$ for $1 \leq j \leq k$. The narrow width of this permissible range for $|q(\alpha)\alpha_j - a_j(\alpha)|$ and the coprimality condition ensure that $q(\alpha)$ and $\mathbf{a}(\alpha)$ are well-defined. Let

$$\beta_j(\alpha) = \alpha_j - \frac{a_j(\alpha)}{q(\alpha)}, \quad (1 \leq j \leq k),$$

and

$$d(\alpha) = (q(\alpha), a_2(\alpha), \dots, a_k(\alpha)) \ll 1.$$

Note that for sufficiently large P, we have $P^{-\zeta} \leq (2k^2)^{-1}$. Recalling (3.4.2), we apply Lemma 3.4.5 to conclude that

$$\begin{aligned} f_{\theta_3}(\alpha) &= \sum_{x=1}^{P} e(\alpha(x-\theta_3)^k) \\ &= \sum_{x=1}^{P} e(\alpha_k x^k + \ldots + \alpha_1 x + \alpha_0) \\ &\ll q(\alpha)^{-1} \left|S(q(\alpha), \mathbf{a}(\alpha))I(\boldsymbol{\beta}(\alpha))\right| + q(\alpha)^{1-1/k+\epsilon}. \end{aligned}$$

We now use [56, Theorems 7.1 and 7.3] to provide estimates for $S(q(\alpha), \mathbf{a}(\alpha))$ and $I(\boldsymbol{\beta}(\alpha))$. We have

$$S(q(\alpha), \mathbf{a}(\alpha)) \ll q(\alpha)^{1-1/k+\epsilon},$$

and

$$I(\boldsymbol{\beta}(\alpha)) \ll P(1 + |\beta_1(\alpha)|\, P + \ldots + |\beta_k(\alpha)|\, P^k)^{-1/k}.$$

Hence we see that

$$\begin{aligned} f_{\theta_3}(\alpha) &\ll q(\alpha)^{-1/k+\epsilon} P(1 + \ldots + |\beta_k(\alpha)|\, P^k)^{-1/k} + q(\alpha)^{1-1/k+\epsilon} \\ &\ll q(\alpha)^{-1/k+\epsilon} P(1 + |\beta_k(\alpha)|\, P^k)^{-1/k}. \end{aligned} \tag{3.4.3}$$

We now use this result to bound the integral that we are interested in.

Lemma 3.4.6. *For $u > 2k$, we have*

$$\int_{\mathfrak{B}_v} |f_{\theta_3}(\alpha)|^u \, d\alpha \ll P^{u-k}.$$

Proof. By the above definitions, we note that if $q(\alpha) = q(\alpha')$, $a_k(\alpha) = a_k(\alpha')$ and $\beta_k(\alpha) = \beta_k(\alpha')$, then in fact $\alpha = \alpha'$. Using (3.4.3), we therefore have

$$\begin{aligned}\int_{\mathfrak{B}_v} |f_{\theta_3}(\alpha)|^u \, d\alpha &\ll \int_{\mathfrak{B}_v} (q(\alpha)^{-1/k+\epsilon} P(1 + |\beta_k(\alpha)| \, P^k)^{-1/k})^u \, d\alpha \\ &\ll P^u \sum_{1 \le q \le P^{1-\zeta}} \sum_{a_k=1}^{q} q^{-u/k+\epsilon} \int_{|\beta_k| \le P^{1-k-\zeta}} (1 + |\beta_k| \, P^k)^{-u/k} \, d\beta_k,\end{aligned}$$

and so

$$\int_{\mathfrak{B}_v} |f_{\theta_3}(\alpha)|^u \, d\alpha \ll P^u J \sum_{1 \le q \le P^{1-\zeta}} q^{1-u/k+\epsilon},$$

where, just as in [15, Corollary 2.4], we have

$$J = \int_0^\infty (1 + \beta P^k)^{-u/k} \, d\beta \ll P^{-k}.$$

Consequently, since $u/k > 2$ and ϵ is small, we see that

$$\begin{aligned}\int_{\mathfrak{B}_v} |f_{\theta_3}(\alpha)|^u \, d\alpha &\ll P^{u-k} \sum_{1 \le q \le P^{1-\zeta}} q^{1-u/k+\epsilon} \\ &\ll P^{u-k}.\end{aligned}$$

□

Lemma 3.4.7. *For $s > 2k + 2$ and for v with $\mathfrak{B}_v \subset \mathfrak{m}$, we have*

$$\int_{\mathfrak{B}_v} \left| f_{\theta_1}(\alpha) f_{\theta_2}(\alpha) f_{\theta_3}(\alpha)^{s-2} \right| \, d\alpha \ll P^{s-k} \, T(P)^{-1}.$$

Proof. Using (3.2.3), we have

$$\begin{aligned}\int_{\mathfrak{B}_v} \left|f_{\theta_1}(\alpha)f_{\theta_2}(\alpha)f_{\theta_3}(\alpha)^{s-2}\right| \, d\alpha &\ll \sup_{\alpha\in\mathfrak{B}_v} |f_{\theta_1}(\alpha)f_{\theta_2}(\alpha)| \int_{\mathfrak{B}_v} |f_{\theta_3}(\alpha)|^{s-2} \, d\alpha \\ &\ll P^2\, T(P)^{-1} \int_{\mathfrak{B}_v} |f_{\theta_3}(\alpha)|^{s-2} \, d\alpha.\end{aligned}$$

Since $s > 2k+2$, we may apply Lemma 3.4.6 with $u = s-2$ to obtain

$$\begin{aligned}\int_{\mathfrak{B}_v} \left|f_{\theta_1}(\alpha)f_{\theta_2}(\alpha)f_{\theta_3}(\alpha)^{s-2}\right| \, d\alpha &\ll P^2\, T(P)^{-1} P^{s-2-k} \\ &\ll P^{s-k}\, T(P)^{-1}.\end{aligned}$$

□

We now combine the minor and trivial arc estimates to deduce the required result for the whole of $\mathfrak{B}$.

Lemma 3.4.8. *We have*

$$\int_{\mathfrak{B}} \left|f_{\theta_1}(\alpha)f_{\theta_2}(\alpha)f_{\theta_3}(\alpha)^{s-2}K_{\pm}(\alpha)\right| \, d\alpha = o(P^{s-k}).$$

Proof. As in Lemma 3.4.3, we split the integral over $\mathfrak{B}$ into integrals over $\mathfrak{B}_v$, distinguishing between those intervals contained in $\mathfrak{m}$, and those contained in (or intersecting) $\mathfrak{t}$. Using Lemma 3.4.7 and (3.2.6), and writing $\omega = P^{\xi-k}$ and $z = T(P) - \omega - 1$ for brevity, we have

$$\begin{aligned}&\sum_{0\le v\le z} \int_{\mathfrak{B}_{v+\omega}} \left|f_{\theta_1}(\alpha)f_{\theta_2}(\alpha)f_{\theta_3}(\alpha)^{s-2}K_{\pm}(\alpha)\right| \, d\alpha \\ &\qquad \ll P^{s-k}\, T(P)^{-1} + \sum_{1\le v\le z} \frac{L(P)}{(v+\omega)^2} \int_{\mathfrak{B}_{v+\omega}} \left|f_{\theta_1}(\alpha)f_{\theta_2}(\alpha)f_{\theta_3}(\alpha)^{s-2}\right| \, d\alpha \\ &\qquad \ll P^{s-k}\, T(P)^{-1} + P^{s-k}\, \frac{L(P)}{T(P)} \sum_{1\le v\le z} \frac{1}{(v+\omega)^2} \\ &\qquad = o(P^{s-k}),\end{aligned}$$

while, by Lemma 3.4.6,

$$\begin{aligned}\sum_{v=-1}^{\infty}\int_{\mathfrak{B}_{v+T(P)}}\left|f_{\theta_1}(\alpha)f_{\theta_2}(\alpha)f_{\theta_3}(\alpha)^{s-2}K_{\pm}(\alpha)\right|\,d\alpha &\ll \sum_{v=-1}^{\infty}\frac{L(P)}{(v+T(P))^2}P^2P^{s-2-k}\\ &\ll \frac{L(P)}{T(P)-1}P^{s-k}\\ &= o(P^{s-k}).\end{aligned}$$

Combining the above sums, we achieve the stated result for the whole of $\mathfrak{B}$. □

We now summarise our conclusion for the whole of the Davenport–Heilbronn minor and trivial arcs in another lemma.

Lemma 3.4.9. *For any natural number $s \geq s_0(k)$, we have*

$$\int_{\mathfrak{m}\cup\mathfrak{t}} f_{\boldsymbol{\theta}}(\alpha)e(-\tau\alpha)K_{\pm}(\alpha)\,d\alpha = o(P^{s-k}).$$

Assuming the shifted Hua's lemma, the same result holds whenever $s \geq s_1(k)$.

Proof. By symmetry, the above results hold equally well when θ_3 is replaced by any other θ_i with $i \geq 4$. Applying Hölder's inequality,

$$\begin{aligned}\int_{\mathfrak{N}} f_{\boldsymbol{\theta}}(\alpha)e(-\tau\alpha)K_{\pm}(\alpha)\,d\alpha &\ll \prod_{i=3}^{s}\left(\int_{\mathfrak{N}}\left|f_{\theta_1}(\alpha)f_{\theta_2}(\alpha)f_{\theta_i}(\alpha)^{s-2}K_{\pm}(\alpha)\right|\,d\alpha\right)^{1/(s-2)}\\ &= o(P^{s-k}),\end{aligned}$$

by Lemmata 3.4.3 and 3.4.8. By Corollary 3.3.6, and using the dissection $\mathfrak{m}\cup\mathfrak{t} = \mathfrak{N}\cup\mathfrak{n}$, we achieve the desired conclusion. □

3.5 The major arc

On the major arc

$$\mathfrak{M} = \{\alpha \in \mathbb{R} : |\alpha| < P^{\xi-k}\},$$

we use a result of Chow, noting that it requires only that the number of variables be greater than k.

Lemma 3.5.1. *We have*

$$\int_{\mathfrak{M}} f_{\boldsymbol{\theta}}(\alpha)e(-\tau\alpha)K_{\pm}(\alpha)\,d\alpha = 2\eta\Gamma(1+1/k)^s\Gamma(s/k)^{-1}P^{s-k} + o(P^{s-k}).$$

Proof. This is [15, equation (3.28)]. □

Combining Lemmata 3.4.9 and 3.5.1 with (3.2.4), we obtain the conclusion of Theorem 3.1.1. Assuming the shifted Hua's lemma, we would achieve the same result whenever $s \geq s_1(k)$. In particular, this would provide a further improvement when $k = 10$ or $k \geq 12$.

3.6 Almost equal summands

One variant of the above problem is to consider solutions of (3.1.1) subject to the additional condition

$$\left|x_i - (\tau/s)^{1/k}\right| < y(\tau), \quad (1 \leq i \leq s),$$

for some function $y(\tau)$. This is a natural direction to consider, since it has been studied in the classical case of Waring's problem, and as mentioned in Section 1.2, Wright proved in [72] that we cannot guarantee solutions in the classical case if the function y is too small.

In this section, we show that (a slight strengthening of) Wright's result remains true in the shifted case, despite the fact that we are no longer dealing with a purely integer problem. Specifically, we prove the following.

Theorem 3.6.1. *Let $s, k \geq 2$ be natural numbers. Fix $\boldsymbol{\theta} = (\theta_1, \dots, \theta_s) \in (0,1)^s$, and let $c, c' > 0$ be suitably small constants which may depend on s, k and $\boldsymbol{\theta}$. There exist arbitrarily large values of $\tau \in \mathbb{R}$ which cannot be approximated in the form (3.1.1), with $0 < \eta < c\tau^{1-2/k}$, subject to the additional condition that $\left|x_i - (\tau/s)^{1/k}\right| < c'\tau^{1/2k}$ for $1 \leq i \leq s$.*

Proof. This follows the structure of Wright's proof in [72], with minor adjustments to take into account the shifts present in our problem. As such, for $m \in \mathbb{N}$, we let $\tau_m = sm^k + km^{k-1}(s - \sum_{i=1}^{s}\theta_i)$, and we note that $\tau_m \to \infty$ as $m \to \infty$. Throughout the proof, we allow $c_1, c_2, \dots$ to denote positive

constants which do not depend on m, although they may depend on the fixed values of $s, k, \boldsymbol{\theta}, c$ and c'. We also note that $\eta < c\tau^{1-2/k}$ implies that $\eta \ll m^{k-2}$.

Suppose τ_m satisfies (3.1.1) with $0 < \eta < c\tau_m^{1-2/k}$ and

$$\left|x_i - (\tau_m/s)^{1/k}\right| < c'\tau_m^{1/2k}$$

for $1 \le i \le s$. We write $x_i = m + a_i$, and observe that

$$\begin{aligned} m^{k-1}\left|a_i\right| &= m^{k-1}\left|x_i - m\right| \\ &\le m^{k-1}\Big(\left|x_i - (\tau_m/s)^{1/k}\right| + \left|(\tau_m/s)^{1/k} - m\right|\Big) \\ &\le c'm^{k-1}\tau_m^{1/2k} + \left|\tau_m/s - m^k\right|. \end{aligned}$$

Using the definition of τ_m, we obtain

$$m^{k-1}\left|a_i\right| \le c_1 m^{k-1}m^{1/2} + km^{k-1}(1 - s^{-1}\sum_{i=1}^{s}\theta_i),$$

and therefore $|a_i| \le c_2 m^{1/2}$ for $1 \le i \le s$.

Expanding (3.1.1), we see that

$$\begin{aligned} \eta &> \left|\sum_{i=1}^{s}(x_i - \theta_i)^k - \tau_m\right| \\ &= \left|\sum_{i=1}^{s}(m + a_i - \theta_i)^k - \Big(sm^k + km^{k-1}(s - \sum_{i=1}^{s}\theta_i)\Big)\right| \qquad (3.6.1) \\ &\ge km^{k-1}\left|s - \sum_{i=1}^{s}a_i\right| - \left|\sum_{j=2}^{k}\binom{k}{j}m^{k-j}\sum_{i=1}^{s}(a_i - \theta_i)^j\right|. \end{aligned}$$

Rearranging, this gives

$$\begin{aligned} \left|s - \sum_{i=1}^{s}a_i\right| &< \eta k^{-1}m^{1-k} + \left|\sum_{j=2}^{k}\binom{k}{j}k^{-1}m^{1-j}\sum_{i=1}^{s}(a_i - \theta_i)^j\right| \\ &\le \eta k^{-1}m^{1-k} + \sum_{j=2}^{k}\binom{k}{j}k^{-1}m^{1-j}s(c_3m^{1/2})^j \\ &\le c_4. \end{aligned}$$

By choosing our original c, c' to be sufficiently small, we may conclude that

$c_4 \leq 1$, which implies that $\sum_{i=1}^{s} a_i = s$. Substituting this back into (3.6.1), when $k = 2$ we obtain

$$\eta > \binom{k}{2} m^{k-2} \sum_{i=1}^{s} (a_i - \theta_i)^2,$$

and consequently

$$\sum_{i=1}^{s} (a_i - \theta_i)^2 < c_5,$$

which is a contradiction if we choose c, c' sufficiently small, since we have $\sum_{i=1}^{s} (a_i - \theta_i)^2 \gg 1$.

When $k \geq 3$, we obtain

$$\begin{aligned} \eta &> \left| \sum_{j=2}^{k} \binom{k}{j} m^{k-j} \sum_{i=1}^{s} (a_i - \theta_i)^j \right| \\ &\geq \binom{k}{2} m^{k-2} \sum_{i=1}^{s} (a_i - \theta_i)^2 - \left| \sum_{j=3}^{k} \binom{k}{j} m^{k-j} \sum_{i=1}^{s} (a_i - \theta_i)^j \right|. \end{aligned}$$

Consequently,

$$\begin{aligned} \binom{k}{2} m^{k-2} \sum_{i=1}^{s} (a_i - \theta_i)^2 &< \eta + \sum_{j=3}^{k} \binom{k}{j} m^{k-j} \sum_{i=1}^{s} |a_i - \theta_i|^j \\ &\leq \eta + \sum_{j=3}^{k} \binom{k}{j} m^{k-j} (c_3 m^{1/2})^{j-2} \sum_{i=1}^{s} (a_i - \theta_i)^2 \leq \eta + c_6 m^{k-5/2} \sum_{i=1}^{s} (a_i - \theta_i)^2, \end{aligned}$$

and so

$$\sum_{i=1}^{s} (a_i - \theta_i)^2 < c_7 + c_8 m^{-1/2} \sum_{i=1}^{s} (a_i - \theta_i)^2,$$

which is again a contradiction when m is large.

We conclude that for all sufficiently large m, it is impossible to approximate τ_m in the manner claimed. This completes the proof. □

Corollary 3.6.2. *For $s, k \geq 2$ natural numbers, $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_s) \in (0,1)^s$, and suitably small constants $C, C' > 0$, there exist arbitrarily wide gaps between real*

numbers τ for which the system

$$\begin{aligned} \left|(x_1-\theta_1)^k+\ldots+(x_s-\theta_s)^k-\tau\right| &< C\tau^{1-2/k} \\ \left|x_i-(\tau/s)^{1/k}\right| &< C'\tau^{1/2k}, \quad (1\le i\le s) \end{aligned} \tag{3.6.2}$$

has a solution in natural numbers $x_1,\ldots,x_s$.

Proof. By Theorem 3.6.1, we fix $\tau_0\in\mathbb{R}$ such that there is no solution in natural numbers $x_1,\ldots,x_s$ to $\left|(x_1-\theta_1)^k+\ldots+(x_s-\theta_s)^k-\tau_0\right| < c\tau_0^{1-2/k}$ with $\left|x_i-(\tau_0/s)^{1/k}\right| < c'\tau_0^{1/2k}$ for $1\le i\le s$.

Let $0<\delta\le C_0\tau_0^{1-2/k}$ for some $C_0>0$, and let $\tau\in[\tau_0-\delta,\tau_0+\delta]$. Let $C,C'>0$ be suitably small constants depending on c,c' and C_0 to be chosen later, and suppose that $x_1\ldots,x_s\in\mathbb{N}$ are such that (3.6.2) is satisfied.

We have

$$\begin{aligned} \left|(\tau/s)^{1/k}-(\tau_0/s)^{1/k}\right| &\le s^{-1/k}\left|(\tau_0-\delta)^{1/k}-\tau_0^{1/k}\right| \\ &\le C_1\delta\tau_0^{1/k-1}, \end{aligned}$$

and consequently

$$\begin{aligned} \left|x_i-(\tau_0/s)^{1/k}\right| &\le \left|x_i-(\tau/s)^{1/k}\right|+\left|(\tau/s)^{1/k}-(\tau_0/s)^{1/k}\right| \\ &< C'\tau^{1/2k}+C_1\delta\tau_0^{1/k-1} \\ &\le C'(\tau_0+\delta)^{1/2k}+C_1C_0\tau_0^{-1/k} \\ &\le C_2\tau_0^{1/2k}. \end{aligned}$$

We also see that

$$\begin{aligned} \left|\sum_{i=1}^s(x_i-\theta_i)^k-\tau_0\right| &\le \left|\sum_{i=1}^s(x_i-\theta_i)^k-\tau\right|+|\tau-\tau_0| \\ &< C\tau^{1-2/k}+\delta \\ &\le C(\tau_0+\delta)^{1-2/k}+C_0\tau_0^{1-2/k} \\ &\le C_3\tau_0^{1-2/k}. \end{aligned}$$

Choosing C_0,C,C' small enough to ensure that $C_2\le c'$ and $C_3\le c$ gives a contradiction to our original choice of τ_0. Consequently, there is no solution to (3.6.2) in an interval of radius $\asymp\tau_0^{1-2/k}$ around τ_0. □

Chapter 4

Efficient Congruencing with Ellipsephic Sets: the quadratic case

The work in this chapter is based on the author's upcoming paper [6].

4.1 Introduction

In this chapter, as discussed in Section 1.3, we investigate a variant of Vinogradov's mean value theorem in which the variables are restricted to certain thin subsets of the natural numbers defined by digital restrictions. We recall some definitions. For a subset $A \subset \mathbb{N}_0$, and a prime p, we let the associated *ellipsephic set* be

$$\mathcal{E} = \mathcal{E}_p^A = \{n \in \mathbb{N} \mid n = \textstyle\sum_i a_i p^i, a_i \in A \cap [0,p) \text{ for all } i\}.$$

We work in the case $2 \leq \#(A \cap [0,p)) \leq p-1$, and observe, writing $\mathcal{E}(X) = \mathcal{E} \cap [1,X]$, that

$$\lim_{X \to \infty} \frac{\#\mathcal{E}(X)}{X} = 0.$$

For $t \geq 2$ an integer, and $\delta > 0$ a real number, we call a set $A \subset \mathbb{N}_0$ an $E_t(\delta)$*-set* if it has the property that

$$\#\{(a_1, \dots, a_t) \in A^t \mid a_1 + \dots + a_t = n\} \ll n^{\delta},$$

and an E_t^**-set* if it has this property for every $\delta > 0$. We refer to a set $\mathcal{E}$ as a (p, t, δ)*-ellipsephic set* if $\mathcal{E} = \mathcal{E}_p^A$ and A is an $E_t(\delta)$-set, and as a $(p, t)^*$*-ellipsephic set* if A is an E_t^*-set.

As mentioned in Section 1.3, the set of squares forms a key motivating example for this work, being an E_2^*-set. In other words, the number of representations of a natural number n in the form a^2+b^2 is $O(n^\epsilon)$—see [42, Hilfssatz 14]—and in discussing our main results below, we present the special case of integers with square digits as a corollary.

We also recall that, for a sequence $\mathfrak{a} = (\mathfrak{a}_x)_{x\in\mathcal{E}}$ of complex weights, we are interested in the quantity

$$J_s(X) = J_{s,2}(X; \mathfrak{a}) = \oint \Big| \sum_{x\in\mathcal{E}(X)} \mathfrak{a}_x e(\alpha_1 x + \alpha_2 x^2)\Big|^{2s} d\boldsymbol{\alpha},$$

which counts the solutions, in positive integers $x_i, y_i \in \mathcal{E}(X)$, to the system

$$x_1^j + \ldots + x_s^j = y_1^j + \ldots + y_s^j, \quad (1 \le j \le 2), \tag{4.1.1}$$

where each solution is counted with weight $\mathfrak{a}_{\boldsymbol{x}}\overline{\mathfrak{a}_{\boldsymbol{y}}} = \mathfrak{a}_{x_1}\ldots\mathfrak{a}_{x_s}\overline{\mathfrak{a}_{y_1}\ldots\mathfrak{a}_{y_s}}$. The main theorem of this chapter is Theorem 1.3.1, which we restate here, and which provides the following upper bound for $J_s(X)$.

Theorem 4.1.1. *For $t \ge 2$ an integer, $\delta > 0$ a real number, and p a suitably large prime, let $\mathcal{E}$ be a (p, t, δ)-ellipsephic set and let $Y = \#\mathcal{E}(X)$. Then for $s \ge 3t$, we have*

$$J_s(X) \ll Y^{s-3t}X^{3\delta+\epsilon}\bigg(\sum_{x\in\mathcal{E}(X)} |\mathfrak{a}_x|^2\bigg)^s.$$

If $\mathcal{E}$ is a $(p, t)^*$-ellipsephic set, we have

$$J_s(X) \ll Y^{s-3t}X^{\epsilon}\bigg(\sum_{x\in\mathcal{E}(X)} |\mathfrak{a}_x|^2\bigg)^s.$$

Corollary 4.1.2. *The unweighted number of solutions to (4.1.1) with $x_i, y_i \in \mathcal{E}(X)$ for all i is $O(Y^{2s-3t}X^{3\delta+\epsilon})$.*

Proof. This is the case where $\mathfrak{a}_x = 1$ for all $x \in \mathcal{E}$. □

Corollary 4.1.3. *In the case where $\mathcal{E} = \mathcal{E}_p^{A_2}$, with $A_2 = \{n^2 \mid n \in \mathbb{N}_0\}$, we*

have that for $s \geq 6$*,*

$$J_s(X) \ll Y^{s-6}X^{\epsilon}\bigg(\sum_{x\in\mathcal{E}(X)}|\mathfrak{a}_x|^2\bigg)^s,$$

and in the unweighted case,

$$J_s(X) \ll Y^{2s-6}X^{\epsilon} \ll Y^{2s-6+\epsilon}.$$

At the critical case $s = 6$*, this yields*

$$J_6(X) \ll Y^{6+\epsilon}.$$

Proof. As mentioned above, A_2 is an E_2^*-set, and the result follows from Theorem 4.1.1. □

As mentioned in Section 1.4.2, Theorem 4.1.1 has potential future applications to a number of other Diophantine problems, most notably Waring's problem, in which we attempt to write all natural numbers as sums of a bounded number of squares of ellipsephic integers. As another corollary of Theorem 4.1.1, we provide a lower bound on the number of integers representable in the form required by Waring's problem. We would expect to need the set $\#\mathcal{E}(X)$ to be sufficiently large to give any chance of being able to represent a significant proportion of the integers up to X, and as such we incorporate this as an extra condition in the below result.

Let $N_s(X) = N_{s,2}^{\mathcal{E}}(X)$ be the number of integers n with $1 \leq n \leq X$ which have a representation as a sum of s squares of integers from $\mathcal{E}$.

Corollary 4.1.4. *For* $t \geq 2$ *an integer and* p *a suitably large prime, let* $\mathcal{E}$ *be a* (p, t, δ)*-ellipsephic set for some* $\delta > 0$*. Assume that* $Y = \#\mathcal{E}(X) \gg X^{1/t}$*. Then for* $s \geq 3t$ *we have*

$$N_s(X) \gg X^{1-3\delta/2-\epsilon}.$$

In the case where $\mathcal{E}$ *is a* $(p, t)^*$*-ellipsephic set, we have* $N_s(X) \gg X^{1-\epsilon}$*.*

Proof. Using Cauchy's inequality, and writing $R(n) = R_{s,2}^{\mathcal{E}}(n)$ for the number of representations of an integer n as a sum of s squares of integers from $\mathcal{E}$, we

have

$$\Big(\sum_{1\le n\le X} R(n)\Big)^2 \le \Big(\sum_{\substack{1\le n\le X\\ R(n)>0}} 1\Big)\Big(\sum_{1\le n\le X} R(n)^2\Big)$$
$$= N_s(X)\Big(\sum_{1\le n\le X} R(n)^2\Big). \tag{4.1.2}$$

We note that

$$\Big(\sum_{1\le n\le X} R(n)\Big)^2 \ge \big(\#\mathcal{E}(\sqrt{X/s})^s\big)^2$$
$$\gg Y^s,$$

and, using Theorem 4.1.1, that

$$\Big(\sum_{1\le n\le X} R(n)^2\Big) \ll X^{1/2}J_s(X^{1/2})$$
$$\ll X^{1/2}(Y^{1/2})^{2s-3t}(X^{1/2})^{3\delta+\epsilon}$$
$$= Y^{s-3t/2}(X^{1/2})^{1+3\delta+\epsilon}.$$

Combining these bounds with (4.1.2), we see that

$$N_s(X) \gg Y^{3t/2}(X^{-1/2})^{1+3\delta+\epsilon},$$

and our additional assumption on the size of Y allows us to conclude that

$$N_s(X) \gg X^{1-3\delta/2-\epsilon},$$

as required. □

Note that the special case with the conclusion that $N_s(X) \gg X^{1-\epsilon}$ applies to the aforementioned example of integers with square digits.

The proof of Theorem 4.1.1 uses a version of Wooley's efficient congruencing method which we outline briefly here. We begin by postulating that $J_s(X)$ is significantly larger than the bound asserted in Theorem 4.1.1, and proceed by aiming to derive a contradiction. We partition our variables into congruence classes modulo powers of the base p, and apply Hölder's inequality to restrict

our variables to lie in certain common congruence classes. The binomial theorem allows us to convert our equations into congruences featuring a subset of our variables, and using their ellipsephic nature and the $E_t(\delta)$ property, we can 'lift' solutions to these congruences, at a cost, to diagonal solutions in which each pair of variables is mutually congruent modulo the relevant power of p. Iterating this process, we strengthen the congruences satisfied by these variables—this may be viewed as a 'p-adic concentration' argument, since our variables become closer p-adically. By iterating sufficiently many times, we find that our initial assumption on $J_s(X)$ is untenable, which leads us to a contradiction. In Section 4.2, we provide a series of preliminary results which form the basis of our iteration process, and in Section 4.3 we complete the proof of Theorem 4.1.1.

4.2 Preliminaries

We first observe that the case $s > 3t$ of Theorem 4.1.1 follows directly from the case $s = 3t$, and so we work only in this latter case throughout. We also note that it suffices to prove Theorem 4.1.1 for X a power of p because, for $p^{C-1} < X < p^C$, we then have

$$J(X) \ll J(p^C) \ll (pX)^{3\delta+\epsilon}\bigg(\sum_{x\in\mathcal{E}(p^C)}|\mathfrak{a}_x|^2\bigg)^s$$

for any choice of $\mathfrak{a}$, and so

$$J(X) \ll X^{3\delta+\epsilon}\bigg(\sum_{x\in\mathcal{E}(X)}|\mathfrak{a}_x|^2\bigg)^s.$$

We apply the following normalisation. Let

$$\rho_0 = \bigg(\sum_{x\in\mathcal{E}(X)}|\mathfrak{a}_x|^2\bigg)^{1/2}.$$

For $\boldsymbol{\alpha} \in [0,1]^2$, we let

$$f(\boldsymbol{\alpha}) = f(\boldsymbol{\alpha};\mathfrak{a}) = \rho_0^{-1}\sum_{x\in\mathcal{E}(X)}\mathfrak{a}_x e(\alpha_1 x + \alpha_2 x^2),$$

and define the normalised mean value

$$\mathfrak{J}(X) = \mathfrak{J}_{s,2}(X;\mathfrak{a}) = \oint \big|f(\boldsymbol{\alpha};\mathfrak{a})\big|^{2s}\, d\boldsymbol{\alpha} = \rho_0^{-2s} J(X).$$

Note that this normalisation allows us to assume that $|\mathfrak{a}_x| \le 1$ for all $x \in \mathcal{E}$. We may also restrict ourselves to the situation in which our weights are real and non-negative, as follows. Let $\mathfrak{a}_x = \mathfrak{b}_x^+ - \mathfrak{b}_x^- + i\mathfrak{c}_x^+ - i\mathfrak{c}_x^-$, where $\mathfrak{b}_x^+, \mathfrak{b}_x^-, \mathfrak{c}_x^+$ and $\mathfrak{c}_x^-$ are non-negative real numbers, with at most one of $\mathfrak{b}_x^+$ and $\mathfrak{b}_x^-$ non-zero, and at most one of $\mathfrak{c}_x^+$ and $\mathfrak{c}_x^-$ non-zero. Write

$$g_1(\boldsymbol{\alpha}) = \sum_{x\in\mathcal{E}(X)} \mathfrak{b}_x^+ e(\alpha_1 x + \alpha_2 x^2), \qquad g_2(\boldsymbol{\alpha}) = \sum_{x\in\mathcal{E}(X)} \mathfrak{b}_x^- e(\alpha_1 x + \alpha_2 x^2),$$
$$g_3(\boldsymbol{\alpha}) = \sum_{x\in\mathcal{E}(X)} \mathfrak{c}_x^+ e(\alpha_1 x + \alpha_2 x^2), \qquad g_4(\boldsymbol{\alpha}) = \sum_{x\in\mathcal{E}(X)} \mathfrak{c}_x^- e(\alpha_1 x + \alpha_2 x^2),$$

and observe that

$$\sum_{x\in\mathcal{E}(X)} \mathfrak{a}_x e(\alpha_1 x + \alpha_2 x^2) = g_1(\boldsymbol{\alpha}) - g_2(\boldsymbol{\alpha}) + ig_3(\boldsymbol{\alpha}) - ig_4(\boldsymbol{\alpha}) = \sum_{j=1}^{4} \epsilon_j g_j(\boldsymbol{\alpha}),$$

where we have chosen $\epsilon_j \in \{\pm 1, \pm i\}$ appropriately. By Hölder's inequality, we now split up the integrals we are interested in into the parts corresponding to each of these weights, to see that

$$\oint \Big| \sum_{x\in\mathcal{E}(X)} \mathfrak{a}_x e(\alpha_1 x + \alpha_2 x^2)\Big|^{2s} d\boldsymbol{\alpha} = \oint \Big| \sum_{j=1}^{4} \epsilon_j g_j(\boldsymbol{\alpha})\Big|^{2s} d\boldsymbol{\alpha} \ll \max_j \oint \big|g_j(\boldsymbol{\alpha})\big|^{2s} d\boldsymbol{\alpha},$$

and that since $|\mathfrak{b}_x^{\pm}|, |\mathfrak{c}_x^{\pm}| \le |\mathfrak{a}_x|$, we obtain the required bounds for general weights from those for real, non-negative weights as claimed. We let

$$\mathbb{D} = \{\mathfrak{a} \mid \mathfrak{a}_x \in [0,1] \text{ for all } x \in \mathcal{E}\},$$

and from now on we work with $\mathfrak{a} \in \mathbb{D}$.

With the above normalisation, we see that an estimate of the desired form

$$J(X) \ll X^{\Delta}\Big(\sum_{x\in\mathcal{E}(X)} |\mathfrak{a}_x|^2 \Big)^s,$$

for some $\Delta > 0$, follows directly from one of the form

$$\mathfrak{J}(X) \ll X^{\Delta}.$$

We define

$$\lambda = \sup_{\mathfrak{a} \in \mathbb{D}} \limsup_{X \to \infty} \frac{\log \mathfrak{J}(X; \mathfrak{a})}{\log X}.$$

An application of the Cauchy–Schwarz inequality gives us the trivial bound $\lambda \leq s$. Taking into account the expected value of λ, we define $\Lambda = \lambda - 3\delta$ for ease of notation.

We introduce a series of interdependent constants which come into play during the proof of Theorem 4.1.1 and the results of this section. Let $\epsilon_0 > 0$, and suppose $\Lambda > \epsilon_0$. This is the assumption which we ultimately contradict in Section 4.3.

Let $n = \lceil 12t/\Lambda \rceil$, which will be the number of iterations of the main process in Section 4.3. Let $\iota = \lambda/2^{2n+3}$, and observe that by the definition of λ, there exists a sequence $(X_m)_{m=1}^{\infty}$ tending to infinity with the property that for some $\mathfrak{a} \in \mathbb{D}$, and for large enough m, we have

$$\mathfrak{J}(X_m; \mathfrak{a}) > X_m^{\lambda - \iota}.$$

Henceforth, we work with a choice of $\mathfrak{a} \in \mathbb{D}$ satisfying this condition. In addition, for any $\mathfrak{b} \in \mathbb{D}$, we have

$$\mathfrak{J}(X; \mathfrak{b}) \ll X^{\lambda + \iota}.$$

Suppose that $X = p^B$, where $B \in \mathbb{N}$ is a large parameter satisfying $B \geq 2^{n+3}$, and also that X is sufficiently large with regards to the sequence (X_m). The proof of our main theorem features ν preliminary steps to handle solutions in which variables are congruent modulo small powers of p, as well as an initialisation step of size p^u, where ν and u are large in some respects, but small in relation to B. Specifically, let $\nu = \lceil B/2^{2n+2} \rceil$ and $u = \lceil B/2^{n+2} \rceil$.

While we would usually expect them to be significantly larger, we note that we certainly have $n \geq 4$ and $u \geq 2$. Additionally, we record two further

bounds which will come into play in Section 4.3. Namely, we have

$$\begin{aligned} 2^n(u+1)+\nu-1 &\le B/4+2^{n+1}+B/2^{2n+2} \\ &\le B/2+B/2^{2n+2}<B, \end{aligned} \tag{4.2.1}$$

and

$$\begin{aligned} 2tu-\nu &\ge 2tB/2^{n+2}-B/2^{2n+2}-1 \\ &\ge (2^{n+1}t-1-2^{n-1})B/2^{2n+2} \\ &> \lambda B/2^{2n+2}=2\iota B. \end{aligned} \tag{4.2.2}$$

Our work is heavily dependent on the partition of our variables into congruence classes modulo various powers of p, and we therefore wish to define the restriction of $f(\boldsymbol{\alpha})$ to such classes. For $a \in \mathbb{N}$ and $\xi \in \mathcal{E}(p^a)$, let

$$\rho_a(\xi)=\Bigg(\sum_{\substack{x\in\mathcal{E}(X)\\ x\equiv\xi \pmod{p^a}}} \mathfrak{a}_x^2\Bigg)^{1/2}$$

and

$$f_a(\boldsymbol{\alpha},\xi)=\rho_a(\xi)^{-1}\sum_{\substack{x\in\mathcal{E}(X)\\ x\equiv\xi \pmod{p^a}}} \mathfrak{a}_x e(\alpha_1 x+\alpha_2 x^2).$$

For convenience, we let $\rho_0(\xi)=\rho_0$ and $f_0(\boldsymbol{\alpha},\xi)=f(\boldsymbol{\alpha})$ for any ξ.

We observe that for any $a \in \mathbb{N}$, we have

$$\sum_{\xi\in\mathcal{E}(p^a)} \rho_a(\xi)^2=\rho_0^2,$$

and more generally, for $a,b \in \mathbb{N}$ with $a \le b$,

$$\sum_{\substack{\xi'\in\mathcal{E}(p^b)\\ \xi'\equiv\xi \pmod{p^a}}} \rho_b(\xi')^2=\rho_a(\xi)^2.$$

We will be interested in the following expressions, the first of which represents the weighted number of solutions to our system of equations in which the variables fall into certain congruence classes modulo powers of p. For $a,b \in \mathbb{N}$,

we let

$$I_{a,b}(\xi,\eta) = \oint |f_a(\boldsymbol{\alpha},\xi)|^{2t}\,|f_b(\boldsymbol{\alpha},\eta)|^{4t}\,d\boldsymbol{\alpha},$$

and $I_{0,0}(\xi,\eta) = \mathfrak{J}(X)$ for any ξ and η.

Next, a weighted sum over the possible values of ξ and η in the above definition will simplify later computations. For $h \in \mathbb{N}$, we define

$$K_{a,b}^{h} = \rho_0^{-4} \sum_{\xi\in\mathcal{E}(p^a)} \sum_{\substack{\eta\in\mathcal{E}(p^b)\\ p^{h-1}\|(\xi-\eta)}} \rho_a(\xi)^2\rho_b(\eta)^2 I_{a,b}(\xi,\eta), \tag{4.2.3}$$

where the notation $p^c\|d$ means that $p^c \mid d$ and $p^{c+1} \nmid d$.

The next two lemmata provide useful upper bounds required for completion of the proof of Theorem 4.1.1.

Lemma 4.2.1. *For $a \in \mathbb{N}$ with $p^a < X$, we have*

$$\oint |f_a(\boldsymbol{\alpha},\xi)|^{6t}\,d\boldsymbol{\alpha} \ll (X/p^a)^{\lambda+\iota}.$$

Proof. The above integral counts solutions to the system

$$\sum_{i=1}^{3t}(x_i^2 - y_i^2) = 0 = \sum_{i=1}^{3t}(x_i - y_i)$$

with $x_i, y_i \in \mathcal{E}(X)$ for $1 \le i \le 3t$ and $\boldsymbol{x} \equiv \boldsymbol{y} \equiv \xi \pmod{p^a}$, where solutions are counted with weight $\mathfrak{a}_{\boldsymbol{x}}\mathfrak{a}_{\boldsymbol{y}}\rho_a(\xi)^{-6t}$. Writing $x_i = p^a z_i + \xi$ and $y_i = p^a w_i + \xi$ for $1 \le i \le 3t$, and defining a new set of weights $\mathfrak{b}_z = \rho_a(\xi)^{-1}\mathfrak{a}_{p^a z+\xi}$, we can reinterpret the above system in the form

$$\sum_{i=1}^{3t}(z_i^2 - w_i^2) = 0 = \sum_{i=1}^{3t}(z_i - w_i)$$

with $z_i, w_i \in \mathcal{E}\big((X-\xi)/p^a\big)$ for $1 \le i \le 3t$ and solutions counted with weight $\mathfrak{b}_{\boldsymbol{z}}\mathfrak{b}_{\boldsymbol{w}}$. By definition, this is $\mathfrak{J}\big((X-\xi)/p^a;\mathfrak{b}\big)$, and consequently we have

$$\begin{aligned}\oint |f_a(\boldsymbol{\alpha},\xi)|^{6t}\,d\boldsymbol{\alpha} &\ll \mathfrak{J}(X/p^a;\mathfrak{b})\\ &\ll (X/p^a)^{\lambda+\iota}.\end{aligned}$$

□

Lemma 4.2.2. *For $a, b, h \in \mathbb{N}$ satisfying $\max\{p^a, p^b\} < X$, we have*

$$I_{a,b}(\xi, \eta) \ll (X/p^{a/3+2b/3})^{\lambda+\iota}$$

and

$$K_{a,b}^h \ll (X/p^{a/3+2b/3})^{\lambda+\iota}.$$

Proof. By definition, and using Hölder's inequality, we see that

$$\begin{aligned} I_{a,b}(\xi, \eta) &= \oint |f_a(\boldsymbol{\alpha}, \xi)|^{2t} |f_b(\boldsymbol{\alpha}, \eta)|^{4t} \, d\boldsymbol{\alpha} \\ &\leq \bigg(\oint |f_a(\boldsymbol{\alpha}, \xi)|^{6t} \, d\boldsymbol{\alpha} \bigg)^{1/3} \bigg(\oint |f_b(\boldsymbol{\alpha}, \eta)|^{6t} \, d\boldsymbol{\alpha} \bigg)^{2/3}. \end{aligned}$$

Applying Lemma 4.2.1, we deduce that

$$\begin{aligned} I_{a,b}(\xi, \eta) &\ll ((X/p^a)^{\lambda+\iota})^{1/3} ((X/p^b)^{\lambda+\iota})^{2/3} \\ &= (X/p^{a/3+2b/3})^{\lambda+\iota}. \end{aligned}$$

The second claim follows from the first, using the definition (4.2.3), which gives

$$\begin{aligned} K_{a,b}^h &\ll (X/p^{a/3+2b/3})^{\lambda+\iota} \rho_0^{-4} \sum_{\xi \in \mathcal{E}(p^a)} \sum_{\substack{\eta \in \mathcal{E}(p^b) \\ p^{h-1} \| (\xi - \eta)}} \rho_a(\xi)^2 \rho_b(\eta)^2 \\ &\leq (X/p^{a/3+2b/3})^{\lambda+\iota} \rho_0^{-4} \bigg(\sum_{\xi \in \mathcal{E}(p^a)} \rho_a(\xi)^2 \bigg) \bigg(\sum_{\eta \in \mathcal{E}(p^b)} \rho_b(\eta)^2 \bigg) \\ &= (X/p^{a/3+2b/3})^{\lambda+\iota}. \end{aligned}$$

□

We want to count solutions to congruences modulo some power p^c in the way that we count solutions to equations, via orthogonality, and as such, we make use of Wooley's notation

$$\oint_{p^c} F(\alpha) \, d\alpha = p^{-c} \sum_{1 \leq u \leq p^c} F(u/p^c),$$

and observe that $\oint_{p^c} |f(\boldsymbol{\alpha})|^{2s} \, d\boldsymbol{\alpha}$ counts the number of solutions to the system

$$\sum_{i=1}^{s} (x_i^j - y_i^j) \equiv 0 \pmod{p^c}, \quad (j = 1, 2)$$

with $\boldsymbol{x}, \boldsymbol{y} \in \mathcal{E}(X)^s$, weighted by $\mathfrak{a}_{\boldsymbol{x}}\mathfrak{a}_{\boldsymbol{y}}\rho_0^{-2s}$.

For $c, d \in \mathbb{N}_0$ with $c \leq d$, weights $\mathfrak{b} = (\mathfrak{b}_x)_{x\in\mathcal{E}}$ with $|\mathfrak{b}_x| \leq 1$ for all $x \in \mathcal{E}$, and $\boldsymbol{z} \in \mathcal{E}(p^c)^t$, we define

$$G_{c,d}(\boldsymbol{z}) = \oint_{p^d} \Big| \sum_{\substack{\boldsymbol{x}\in\mathcal{E}(X)^t \\ \boldsymbol{x}\equiv\boldsymbol{z} \pmod{p^c}}} \mathfrak{b}_{\boldsymbol{x}} e\big(\beta(x_1 + \ldots + x_t)\big)\Big|^2 d\beta.$$

The next lemma provides the key 'lifting' step of the process, in which we make use of the $E_t(\delta)$ property of our digit set to raise the power of p used in our congruences.

Lemma 4.2.3. *We have*

$$G_{c,d}(\boldsymbol{z}) \ll p^{\delta(d-c)} \sum_{\substack{\boldsymbol{u}\in\mathcal{E}(p^d)^t \\ \boldsymbol{u}\equiv\boldsymbol{z} \ (mod\ p^c)}} \Big| \sum_{\substack{\boldsymbol{x}\in\mathcal{E}(X)^t \\ \boldsymbol{x}\equiv\boldsymbol{u} \ (mod\ p^d)}} \mathfrak{b}_{\mathbf{x}} \Big|^2.$$

Proof. We observe that $G_{c,d}(\boldsymbol{z})$ counts solutions to the congruence

$$\sum_{i=1}^{t} x_i \equiv \sum_{i=1}^{t} y_i \pmod{p^d} \tag{4.2.4}$$

with $\boldsymbol{x}, \boldsymbol{y} \in \mathcal{E}(X)^t$ and $\boldsymbol{x} \equiv \boldsymbol{y} \equiv \boldsymbol{z} \pmod{p^c}$, with weight $\mathfrak{b}_{\boldsymbol{x}}\overline{\mathfrak{b}_{\boldsymbol{y}}}$. For $1 \leq i \leq t$, let

$$x_i = z_i + \sum_{r\geq c} x_i^{(r)} p^r$$

and

$$y_i = z_i + \sum_{r\geq c} y_i^{(r)} p^r,$$

with $x_i^{(r)}, y_i^{(r)} \in A_p$ for $1 \leq i \leq t$ and $r \geq c$. We bound the number of solutions to (4.2.4) by considering each base p digit in turn. Let

$$\mathcal{A}_t(h) = \Big\{ \boldsymbol{u} \in A_p^t \;\Big|\; \sum_{i=1}^{t} u_i = h \Big\},$$

and

$$\widetilde{\mathcal{A}}_t(h) = \Big\{ (\boldsymbol{u}, \boldsymbol{v}) \in A_p^{2t} \;\Big|\; \sum_{i=1}^{t} (u_i - v_i) = h \Big\}.$$

Summing the lowest digits which interest us (namely, those corresponding to the p^c term in the base p expansion of our variables), we see that a solution of (4.2.4) satisfies

$$(\boldsymbol{x}^{(c)}, \boldsymbol{y}^{(c)}) \in \widetilde{\mathcal{A}}_t(\lambda_c p)$$

for some $1-t \leq \lambda_c \leq t-1$. Accounting for this carry-over between digits, and moving on to the next highest digits, we then see that

$$(\boldsymbol{x}^{(c+1)}, \boldsymbol{y}^{(c+1)}) \in \widetilde{\mathcal{A}}_t(\lambda_{c+1} p - \lambda_c)$$

for some $1 - t \leq \lambda_{c+1} \leq t - 1$. Continuing this process, and setting $\lambda_{c-1} = 0$ for convenience, we obtain the system

$$(\boldsymbol{x}^{(r)}, \boldsymbol{y}^{(r)}) \in \widetilde{\mathcal{A}}_t(\lambda_r p - \lambda_{r-1}), \quad (c \leq r \leq d-1).$$

For brevity, we use the notation $\underline{\boldsymbol{u}}$ to denote the tuple $(\boldsymbol{u}^{(c)}, \dots, \boldsymbol{u}^{(d-1)})$—this represents a regrouping of our variables by digit—and similarly we use $(\underline{\boldsymbol{u}}, \underline{\boldsymbol{v}})$ for $\Big((\boldsymbol{u}^{(c)}, \boldsymbol{v}^{(c)}), \dots, (\boldsymbol{u}^{(d-1)}, \boldsymbol{v}^{(d-1)})\Big)$. We write

$$\mathcal{A}_t(\boldsymbol{h}) = \left\{ \underline{\boldsymbol{u}} \in A_p^{t(d-c)} \;\middle|\; \boldsymbol{u}^{(r)} \in \mathcal{A}_t(h_r) \text{ for } c \leq r \leq d-1 \right\}$$

and

$$\widetilde{\mathcal{A}}_t(\boldsymbol{h}) = \left\{ (\underline{\boldsymbol{u}}, \underline{\boldsymbol{v}}) \in A_p^{2t(d-c)} \;\middle|\; (\boldsymbol{u}^{(r)}, \boldsymbol{v}^{(r)}) \in \widetilde{\mathcal{A}}_t(h_r) \text{ for } c \leq r \leq d-1 \right\}.$$

By convention, we suppose that for any $\underline{\boldsymbol{u}} = (\boldsymbol{u}^{(c)}, \dots, \boldsymbol{u}^{(d-1)}) \in \mathcal{A}_t(\boldsymbol{h})$, we have $u_i = z_i + \sum_{c \leq r \leq d-1} u_i^{(r)} p^r$, and similarly for $(\underline{\boldsymbol{u}}, \underline{\boldsymbol{v}}) \in \widetilde{\mathcal{A}}_t(\boldsymbol{h})$.

For $\boldsymbol{\lambda} = (\lambda_c, \dots, \lambda_{d-1}) \in \{1-t, \dots, t-1\}^{d-c}$, we write

$$\boldsymbol{\lambda'} = (\lambda_c p - \lambda_{c-1}, \dots, \lambda_{d-1} p - \lambda_{d-2}).$$

We now observe that

$$G_{c,d}(\boldsymbol{z}) = \sum_{\boldsymbol{\lambda} \in \{1-t,\dots,t-1\}^{d-c}} \sum_{(\underline{\boldsymbol{u}}, \underline{\boldsymbol{v}}) \in \widetilde{\mathcal{A}}_t(\boldsymbol{\lambda'})} \sum_{\substack{\boldsymbol{x}, \boldsymbol{y} \in \mathcal{E}(X)^t \\ (\boldsymbol{x}, \boldsymbol{y}) \equiv (\boldsymbol{u}, \boldsymbol{v}) \pmod{p^d}}} \mathfrak{b}_{\mathbf{x}} \overline{\mathfrak{b}_{\mathbf{y}}}.$$

Writing

$$\mathfrak{B}(\boldsymbol{u}) = \sum_{\substack{\boldsymbol{x}\in\mathcal{E}(X)^t\\ \boldsymbol{x}\equiv\boldsymbol{u} \pmod{p^d}}} \mathfrak{b}_{\mathbf{x}}$$

and

$$\phi_{\boldsymbol{u}}(\boldsymbol{\gamma}) = \gamma_c \sum_{i=1}^{t} u_i^{(c)} + \ldots + \gamma_{d-1} \sum_{i=1}^{t} u_i^{(d-1)}$$

for brevity, and encoding the condition $(\underline{\boldsymbol{u}}, \underline{\boldsymbol{v}}) \in \widetilde{\mathcal{A}}_t(\boldsymbol{\lambda'})$ in integral form, we see that

$$\begin{aligned}\sum_{(\underline{\boldsymbol{u}},\underline{\boldsymbol{v}})\in\widetilde{\mathcal{A}}_t(\boldsymbol{\lambda'})} \mathfrak{B}(\boldsymbol{u})\overline{\mathfrak{B}(\boldsymbol{v})} &= \sum_{(\underline{\boldsymbol{u}},\underline{\boldsymbol{v}})\in A_p^{2t(d-c)}} \mathfrak{B}(\boldsymbol{u})\overline{\mathfrak{B}(\boldsymbol{v})} \oint e(\phi_{\boldsymbol{u}}(\boldsymbol{\gamma}) - \phi_{\boldsymbol{v}}(\boldsymbol{\gamma}))e(-\boldsymbol{\gamma}\cdot\boldsymbol{\lambda'})\, d\boldsymbol{\gamma}\\ &= \oint e(-\boldsymbol{\gamma}\cdot\boldsymbol{\lambda'}) \sum_{(\underline{\boldsymbol{u}},\underline{\boldsymbol{v}})\in A_p^{2t(d-c)}} \mathfrak{B}(\boldsymbol{u})\overline{\mathfrak{B}(\boldsymbol{v})}e(\phi_{\boldsymbol{u}}(\boldsymbol{\gamma}) - \phi_{\boldsymbol{v}}(\boldsymbol{\gamma}))\, d\boldsymbol{\gamma}\\ &\le \oint \Big| \sum_{(\underline{\boldsymbol{u}},\underline{\boldsymbol{v}})\in A_p^{2t(d-c)}} \mathfrak{B}(\boldsymbol{u})\overline{\mathfrak{B}(\boldsymbol{v})}e(\phi_{\boldsymbol{u}}(\boldsymbol{\gamma}) - \phi_{\boldsymbol{v}}(\boldsymbol{\gamma}))\Big|\, d\boldsymbol{\gamma}.\end{aligned}$$

The expression on the right-hand side is now independent of our choice of $\boldsymbol{\lambda}$, so we conclude that

$$\begin{aligned}G_{c,d}(\boldsymbol{z}) &\le (2t-1)^{d-c} \oint \Big| \sum_{\underline{\boldsymbol{u}}\in A_p^{t(d-c)}} \mathfrak{B}(\boldsymbol{u})e(\phi_{\boldsymbol{u}}(\boldsymbol{\gamma}))\Big|^2 d\boldsymbol{\gamma}\\ &\ll \sum_{(\underline{\boldsymbol{u}},\underline{\boldsymbol{v}})\in\widetilde{\mathcal{A}}_t(\mathbf{0})} \sum_{\substack{\boldsymbol{x},\boldsymbol{y}\in\mathcal{E}(X)^t\\ (\boldsymbol{x},\boldsymbol{y})\equiv(\boldsymbol{u},\boldsymbol{v}) \pmod{p^d}}} \mathfrak{b}_{\mathbf{x}}\overline{\mathfrak{b}_{\mathbf{y}}}\\ &= \sum_{0\le\boldsymbol{n}\le t(p-1)} \Big| \sum_{\underline{\boldsymbol{u}}\in\mathcal{A}_t(\boldsymbol{n})} \sum_{\substack{\boldsymbol{x}\in\mathcal{E}(X)^t\\ \boldsymbol{x}\equiv\boldsymbol{u} \pmod{p^d}}} \mathfrak{b}_{\mathbf{x}}\Big|^2.\end{aligned}$$

Using Cauchy's inequality, we see that

$$G_{c,d}(\boldsymbol{z}) \ll \sum_{0\le\boldsymbol{n}\le t(p-1)} \Big(\sum_{\underline{\boldsymbol{u}}\in\mathcal{A}_t(\boldsymbol{n})} \Big| \sum_{\substack{\boldsymbol{x}\in\mathcal{E}(X)^t\\ \boldsymbol{x}\equiv\boldsymbol{u} \pmod{p^d}}} \mathfrak{b}_{\mathbf{x}}\Big|^2\Big)\Big(\sum_{\underline{\boldsymbol{u}}\in\mathcal{A}_t(\boldsymbol{n})} 1\Big).$$

From our initial assumption that $\mathcal{E}$ is a (p,t,δ)-ellipsephic set, we know that

for $\boldsymbol{n} = (n_c, \dots, n_{d-1})$ with $0 \le \boldsymbol{n} \le t(p-1)$, we have

$$|\mathcal{A}_t(\boldsymbol{n})| \ll \prod_{r=c}^{d-1} n_r^{\delta} \ll p^{\delta(d-c)},$$

and consequently

$$\begin{aligned} G_{c,d}(\boldsymbol{z}) &\ll p^{\delta(d-c)} \sum_{0 \le \boldsymbol{n} \le t(p-1)} \sum_{\underline{\boldsymbol{u}} \in \mathcal{A}_t(\boldsymbol{n})} \Big| \sum_{\substack{\boldsymbol{x} \in \mathcal{E}(X)^t \\ \boldsymbol{x} \equiv \boldsymbol{u} \pmod{p^d}}} \mathfrak{b}_{\mathbf{x}} \Big|^2 \\ &= p^{\delta(d-c)} \sum_{\substack{\boldsymbol{u} \in \mathcal{E}(p^d)^t \\ \boldsymbol{u} \equiv \boldsymbol{z} \pmod{p^c}}} \Big| \sum_{\substack{\boldsymbol{x} \in \mathcal{E}(X)^t \\ \boldsymbol{x} \equiv \boldsymbol{u} \pmod{p^d}}} \mathfrak{b}_{\mathbf{x}} \Big|^2, \end{aligned}$$

as claimed. □

The next lemma allows us to apply Lemma 4.2.3 as the key ingredient in an iterative process which we use later to complete the proof of Theorem 4.1.1.

Lemma 4.2.4. *For $a, b, h \in \mathbb{N}$ satisfying $h \le a < b \le 2a - h + 1$ and $p^b < X$, we have*

$$K_{a,b}^h \ll p^{\delta(2b-a-h+1)} (X/p^b)^{(\lambda+\iota)/2} (K_{b,2b-h+1}^h)^{1/2}.$$

Proof. We begin by observing that $I_{a,b}(\xi, \eta)$ counts the number of solutions to the system

$$\sum_{i=1}^{t} (x_i^j - y_i^j) = \sum_{l=1}^{2t} (u_l^j - v_l^j), \quad (j = 1, 2)$$

with $x_i, y_i, u_l, v_l \in \mathcal{E}(X)$ for $1 \le i \le t$ and $1 \le l \le 2t$, satisfying $\boldsymbol{x} \equiv \boldsymbol{y} \equiv \xi \pmod{p^a}$ and $\boldsymbol{u} \equiv \boldsymbol{v} \equiv \eta \pmod{p^b}$, and with each solution being counted with weight $\rho_a(\xi)^{-2t} \rho_b(\eta)^{-4t} \mathfrak{a}_{\boldsymbol{x}} \mathfrak{a}_{\boldsymbol{y}} \mathfrak{a}_{\boldsymbol{u}} \mathfrak{a}_{\boldsymbol{v}}$.

Writing $x_i = p^a \tilde{x}_i + \xi$ and $u_l = p^b \tilde{u}_l + \eta$, and similarly for $\boldsymbol{y}$ and $\boldsymbol{v}$, we apply the binomial theorem to see that

$$\sum_{i=1}^{t} ((p^a \tilde{x}_i + \xi - \eta)^j - (p^a \tilde{y}_i + \xi - \eta)^j) = p^{jb} \sum_{l=1}^{2t} (\tilde{u}_l^j - \tilde{v}_l^j), \quad (j = 1, 2),$$

and consequently that we have the congruences

$$\sum_{i=1}^{t} ((p^a \tilde{x}_i + \xi - \eta)^j - (p^a \tilde{y}_i + \xi - \eta)^j) \equiv 0 \pmod{p^{jb}}, \quad (j = 1, 2).$$

In other words, we have

$$\sum_{i=1}^{t}(\tilde{x}_i - \tilde{y}_i) \equiv 0 \pmod{p^{b-a}}, \tag{4.2.5}$$

and

$$p^a \sum_{i=1}^{t}(\tilde{x}_i^2 - \tilde{y}_i^2) + 2(\xi - \eta)\sum_{i=1}^{t}(\tilde{x}_i - \tilde{y}_i) \equiv 0 \pmod{p^{2b-a}}. \tag{4.2.6}$$

We fix the weights appearing in the definition of $G_{c,d}(\boldsymbol{z})$ to be

$$\mathfrak{b}_x = \rho_a(\xi)^{-1}\mathfrak{a}_x e(\alpha_1 x + \alpha_2 x^2).$$

Encoding (4.2.5) as part of our integral, and writing $\boldsymbol{\xi} = (\xi, \dots, \xi)$, we have

$$I_{a,b}(\xi, \eta) = \oint G_{a,b}(\boldsymbol{\xi}) \left|f_b(\boldsymbol{\alpha}, \eta)\right|^{4t} d\boldsymbol{\alpha}.$$

By Lemma 4.2.3, we may conclude that

$$I_{a,b}(\xi, \eta) \ll p^{\delta(b-a)} \sum_{\substack{\boldsymbol{z} \in \mathcal{E}(p^b)^t \\ \boldsymbol{z} \equiv \xi \pmod{p^a}}} \oint \left| \sum_{\substack{\boldsymbol{x} \in \mathcal{E}(X)^t \\ \boldsymbol{x} \equiv \boldsymbol{z} \pmod{p^b}}} \mathfrak{b}_{\boldsymbol{x}} \right|^2 \left|f_b(\boldsymbol{\alpha}, \eta)\right|^{4t} d\boldsymbol{\alpha}.$$

We have therefore introduced, at a cost of $p^{\delta(b-a)}$, the additional condition

$$x_i \equiv y_i \pmod{p^b}, \quad (1 \le i \le t),$$

or equivalently

$$\tilde{x}_i \equiv \tilde{y}_i \pmod{p^{b-a}}, \quad (1 \le i \le t).$$

Substituting this back into (4.2.6), and using the facts that $p^{h-1} \| (\xi - \eta)$ and $h - 1 < a < b$, we see that

$$\sum_{i=1}^{t}(\tilde{x}_i - \tilde{y}_i) \equiv 0 \pmod{p^{b-h+1}}.$$

Encoding this congruence as before, we obtain

$$I_{a,b}(\xi,\eta) \ll p^{\delta(b-a)} \sum_{\substack{\boldsymbol{z}\in\mathcal{E}(p^b)^t\\ \boldsymbol{z}\equiv\xi \pmod{p^a}}} \oint G_{b,a+b-h+1}(\boldsymbol{z}) \left|f_b(\boldsymbol{\alpha},\eta)\right|^{4t} d\boldsymbol{\alpha}.$$

We now apply Lemma 4.2.3 again to see that

$$I_{a,b}(\xi,\eta) \ll p^{\delta(b-h+1)} \sum_{\substack{\boldsymbol{z}\in\mathcal{E}(p^{a+b-h+1})^t\\ \boldsymbol{z}\equiv\xi \pmod{p^a}}} \oint \Big| \sum_{\substack{\boldsymbol{x}\in\mathcal{E}(X)^t\\ \boldsymbol{x}\equiv\boldsymbol{z} \pmod{p^{a+b-h+1}}}} \mathfrak{b}_{\boldsymbol{x}} \Big|^2 \left|f_b(\boldsymbol{\alpha},\eta)\right|^{4t} d\boldsymbol{\alpha},$$

and we have introduced the additional condition

$$\tilde{x}_i \equiv \tilde{y}_i \pmod{p^{b-h+1}}, \quad (1 \le i \le t).$$

Repeating this process, we reach the situation in which

$$\sum_{i=1}^{t} (\tilde{x}_i - \tilde{y}_i) \equiv 0 \pmod{p^{2b-a-h+1}},$$

and a final application of Lemma 4.2.3 allows us to conclude that

$$I_{a,b}(\xi,\eta) \ll p^{\delta(2b-a-h+1)} \sum_{\substack{\boldsymbol{z}\in\mathcal{E}(p^{2b-h+1})^t\\ \boldsymbol{z}\equiv\xi \pmod{p^a}}} \oint \Big| \sum_{\substack{\boldsymbol{x}\in\mathcal{E}(X)^t\\ \boldsymbol{x}\equiv\boldsymbol{z} \pmod{p^{2b-h+1}}}} \mathfrak{b}_{\boldsymbol{x}} \Big|^2 \left|f_b(\boldsymbol{\alpha},\eta)\right|^{4t} d\boldsymbol{\alpha}.$$

Using the definition of $\mathfrak{b}$, and writing $b' = 2b - h + 1$, we deduce that

$$I_{a,b}(\xi,\eta) \ll p^{\delta(b'-a)} \oint \Big(\sum_{\substack{\xi'\in\mathcal{E}(p^{b'})\\ \xi'\equiv\xi \pmod{p^a}}} \rho_a(\xi)^{-2} \rho_{b'}(\xi')^2 \left|f_{b'}(\boldsymbol{\alpha},\xi')\right|^2 \Big)^t \left|f_b(\boldsymbol{\alpha},\eta)\right|^{4t} d\boldsymbol{\alpha}.$$

Applying Hölder's inequality, we see that

$$\begin{aligned} I_{a,b}(\xi,\eta) &\ll p^{\delta(b'-a)} \rho_a(\xi)^{-2} \sum_{\substack{\xi'\in\mathcal{E}(p^{b'})\\ \xi'\equiv\xi \pmod{p^a}}} \rho_{b'}(\xi')^2 \oint \left|f_{b'}(\boldsymbol{\alpha},\xi')\right|^{2t} \left|f_b(\boldsymbol{\alpha},\eta)\right|^{4t} d\boldsymbol{\alpha} \\ &= p^{\delta(b'-a)} \rho_a(\xi)^{-2} \sum_{\substack{\xi'\in\mathcal{E}(p^{b'})\\ \xi'\equiv\xi \pmod{p^a}}} \rho_{b'}(\xi')^2 I_{b',b}(\xi',\eta). \end{aligned}$$

Using Cauchy's inequality and Lemma 4.2.1, we observe that

$$\begin{aligned} I_{b',b}(\xi',\eta) &\le \Big(\oint |f_b(\boldsymbol{\alpha},\eta)|^{2t} \, |f_{b'}(\boldsymbol{\alpha},\xi')|^{4t} \, d\boldsymbol{\alpha} \Big)^{1/2} \Big(\oint |f_b(\boldsymbol{\alpha},\eta)|^{6t} \, d\boldsymbol{\alpha} \Big)^{1/2} \\ &\ll I_{b,b'}(\eta,\xi')^{1/2} (X/p^b)^{(\lambda+\iota)/2}. \end{aligned}$$

Substituting this into (4.2.3), we see that

$$\begin{aligned} K_{a,b}^h &= \rho_0^{-4} \sum_{\xi \in \mathcal{E}(p^a)} \sum_{\substack{\eta \in \mathcal{E}(p^b) \\ p^{h-1} \| (\xi - \eta)}} \rho_a(\xi)^2 \rho_b(\eta)^2 I_{a,b}(\xi,\eta) \\ &\ll p^{\delta(b'-a)} (X/p^b)^{(\lambda+\iota)/2} \rho_0^{-4} \sum_{\eta \in \mathcal{E}(p^b)} \sum_{\substack{\xi' \in \mathcal{E}(p^{b'}) \\ p^{h-1} \| (\xi' - \eta)}} \rho_b(\eta)^2 \rho_{b'}(\xi')^2 I_{b,b'}(\eta,\xi')^{1/2} \\ &\ll p^{\delta(b'-a)} (X/p^b)^{(\lambda+\iota)/2} \rho_0^{-2} \Big(\sum_{\eta \in \mathcal{E}(p^b)} \sum_{\substack{\xi' \in \mathcal{E}(p^{b'}) \\ p^{h-1} \| (\xi' - \eta)}} \rho_b(\eta)^2 \rho_{b'}(\xi')^2 I_{b,b'}(\eta,\xi') \Big)^{1/2} \\ &= p^{\delta(b'-a)} (X/p^b)^{(\lambda+\iota)/2} (K_{b,b'}^h)^{1/2}, \end{aligned}$$

as claimed. □

Finally, the following lemma provides a key step in the iterative process of Section 4.3.

Lemma 4.2.5. *For $h \in \mathbb{N}$, and for $\xi \in \mathcal{E}(p^{h-1})$, we have*

$$I_{h-1,h-1}(\xi,\xi) \ll \rho_{h-1}(\xi)^{-4} \Big(\sum_{\substack{\eta \in \mathcal{E}(p^h) \\ \eta \equiv \xi \ (mod\ p^{h-1})}} \rho_h(\eta)^4 I_{h,h}(\eta,\eta) + |\mathcal{E}(p)|^{2s-2} \, \rho_0^4 \, K_{h,h}^h \Big).$$

Proof. We observe that

$$\begin{aligned} I_{h-1,h-1}(\xi,\xi) &= \oint |f_{h-1}(\boldsymbol{\alpha},\xi)|^{2t} \, |f_{h-1}(\boldsymbol{\alpha},\xi)|^{4t} \, d\boldsymbol{\alpha} \\ &= \oint |f_{h-1}(\boldsymbol{\alpha},\xi)|^{2s} \, d\boldsymbol{\alpha}, \end{aligned}$$

which counts the number of solutions to (4.1.1) with $x_i, y_i \in \mathcal{E}(X)$ for $1 \le i \le s$ and $\boldsymbol{x} \equiv \boldsymbol{y} \equiv \xi \pmod{p^{h-1}}$, each solution being counted with weight $\rho_{h-1}(\xi)^{-2s} \mathfrak{a}_{\boldsymbol{x}} \mathfrak{a}_{\boldsymbol{y}}$.

We partition the solutions based on the congruence classes in which the

variables lie modulo p^h, letting $\mathfrak{J}_h(X,\xi)$ denote the contribution from solutions in which all variables are congruent modulo p^h, and $\mathfrak{J}_h^*(X,\xi)$ the contribution from the remaining solutions, so that

$$I_{h-1,h-1}(\xi,\xi) = \mathfrak{J}_h(X,\xi) + \mathfrak{J}_h^*(X,\xi). \tag{4.2.7}$$

We have

$$\begin{aligned}\mathfrak{J}_h(X,\xi) &= \sum_{\substack{\eta\in\mathcal{E}(p^h)\\ \eta\equiv\xi \pmod{p^{h-1}}}} \rho_{h-1}(\xi)^{-2s}\rho_h(\eta)^{2s} I_{h,h}(\eta,\eta)\\ &\le \rho_{h-1}(\xi)^{-4} \sum_{\substack{\eta\in\mathcal{E}(p^h)\\ \eta\equiv\xi \pmod{p^{h-1}}}} \rho_h(\eta)^4 I_{h,h}(\eta,\eta),\end{aligned} \tag{4.2.8}$$

since $\rho_h(\eta)^2 \le \rho_{h-1}(\xi)^2$.

When estimating $\mathfrak{J}_h^*(X,\xi)$, we may assume, up to a combinatorial factor, that $x_1 \not\equiv x_2 \pmod{p^h}$, and observe that $\mathfrak{J}_h^*(X,\xi)$ is bounded above by at most a constant multiple of

$$\begin{aligned}&\rho_{h-1}(\xi)^{-2} \sum_{\substack{\eta\neq\eta'\in\mathcal{E}(p^h)\\ \eta\equiv\eta'\equiv\xi \pmod{p^{h-1}}}} \rho_h(\eta)\rho_h(\eta') \oint f_h(\boldsymbol{\alpha},\eta) f_h(-\boldsymbol{\alpha},\eta') \left|f_{h-1}(\boldsymbol{\alpha},\xi)\right|^{2s-2} d\boldsymbol{\alpha}\\ &\le \rho_{h-1}(\xi)^{-2} \sum_{\substack{\eta\neq\eta'\in\mathcal{E}(p^h)\\ \eta\equiv\eta'\equiv\xi \pmod{p^{h-1}}}} \rho_h(\eta)\rho_h(\eta') I_{h,h}(\eta,\eta')^{1/2s} I_{h,h}(\eta',\eta)^{1/2s} I_{h-1,h-1}(\xi,\xi)^{1-1/s},\end{aligned}$$

by Hölder's inequality. If $\mathfrak{J}_h^*(X,\xi) = \max\{\mathfrak{J}_h(X,\xi), \mathfrak{J}_h^*(X,\xi)\}$, we have

$$I_{h-1,h-1}(\xi,\xi) \ll \mathfrak{J}_h^*(X,\xi),$$

and may rearrange to obtain

$$\begin{aligned}
I_{h-1,h-1}(\xi,\xi) &\ll \rho_{h-1}(\xi)^{-2s}\Bigg(\sum_{\substack{\eta\neq\eta'\in\mathcal{E}(p^h)\\ \eta\equiv\eta'\equiv\xi \pmod{p^{h-1}}}} \rho_h(\eta)\rho_h(\eta')I_{h,h}(\eta,\eta')^{1/s}\Bigg)^s \\
&\ll \rho_{h-1}(\xi)^{-2s}\Bigg(\sum_{\substack{\eta\neq\eta'\in\mathcal{E}(p^h)\\ \eta\equiv\eta'\equiv\xi \pmod{p^{h-1}}}} \rho_h(\eta)^s\rho_h(\eta')^s I_{h,h}(\eta,\eta')\Bigg)\Bigg(\sum_{\substack{\eta\neq\eta'\in\mathcal{E}(p^h)\\ \eta\equiv\eta'\equiv\xi \pmod{p^{h-1}}}} 1\Bigg)^{s-1} \\
&\ll \rho_{h-1}(\xi)^{-4}\Bigg(\sum_{\substack{\eta,\eta'\in\mathcal{E}(p^h)\\ p^{h-1}\|(\eta-\eta')}} \rho_h(\eta)^2\rho_h(\eta')^2 I_{h,h}(\eta,\eta')\Bigg)|\mathcal{E}(p)|^{2s-2} \\
&= |\mathcal{E}(p)|^{2s-2}\,\rho_{h-1}(\xi)^{-4}\rho_0^4\,K_{h,h}^h. \qquad (4.2.9)
\end{aligned}$$

Substituting (4.2.8) and (4.2.9) into (4.2.7), we deduce that

$$I_{h-1,h-1}(\xi,\xi) \ll \rho_{h-1}(\xi)^{-4}\Bigg(\sum_{\substack{\eta\in\mathcal{E}(p^h)\\ \eta\equiv\xi \pmod{p^{h-1}}}} \rho_h(\eta)^4 I_{h,h}(\eta,\eta) + |\mathcal{E}(p)|^{2s-2}\,\rho_0^4\,K_{h,h}^h\Bigg),$$

as claimed. □

4.3 Proof of Theorem 4.1.1

We first wish to handle those solutions in which all of our variables are congruent modulo some small power of p, since these should contribute neglibly to the total, but would prevent some of the mechanisms of the previous section from working smoothly.

Applying Lemma 4.2.5 twice, we have

$$\begin{aligned}
\mathfrak{J}(X) &\ll \rho_0^{-4}\sum_{\xi\in\mathcal{E}(p)} \rho_1(\xi)^4 I_{1,1}(\xi,\xi) + |\mathcal{E}(p)|^{2s-2}\,K_{1,1}^1 \\
&\ll \rho_0^{-4}\sum_{\xi\in\mathcal{E}(p)}\Bigg(\sum_{\substack{\eta\in\mathcal{E}(p^2)\\ \eta\equiv\xi \pmod{p}}} \rho_2(\eta)^4 I_{2,2}(\eta,\eta) + |\mathcal{E}(p)|^{2s-2}\,\rho_0^4\,K_{2,2}^2\Bigg) + |\mathcal{E}(p)|^{2s-2}\,K_{1,1}^1 \\
&= \rho_0^{-4}\sum_{\eta\in\mathcal{E}(p^2)} \rho_2(\eta)^4 I_{2,2}(\eta,\eta) + |\mathcal{E}(p)|^{2s-1}\,K_{2,2}^2 + |\mathcal{E}(p)|^{2s-2}\,K_{1,1}^1.
\end{aligned}$$

Repeated application of Lemma 4.2.5 therefore yields

$$\mathfrak{J}(X) \ll \rho_0^{-4} \sum_{\omega \in \mathcal{E}(p^\nu)} \rho_\nu(\omega)^4 I_{\nu,\nu}(\omega,\omega) + \sum_{1 \le h \le \nu} |\mathcal{E}(p)|^{2s-3+h} K_{h,h}^h.$$

We have

$$\begin{aligned} I_{\nu,\nu}(\omega,\omega) &= \oint |f_\nu(\boldsymbol{\alpha},\omega)|^{2t} |f_\nu(\boldsymbol{\alpha},\omega)|^{4t} \, d\boldsymbol{\alpha} \\ &= \oint |f_\nu(\boldsymbol{\alpha},\omega)|^{6t} \, d\boldsymbol{\alpha} \ll (X/p^\nu)^{\lambda+\iota}, \end{aligned}$$

by Lemma 4.2.1.

Consequently,

$$\begin{aligned} \rho_0^{-4} \sum_{\omega \in \mathcal{E}(p^\nu)} \rho_\nu(\omega)^4 I_{\nu,\nu}(\omega,\omega) &\ll (X/p^\nu)^{\lambda+\iota} \rho_0^{-4} \sum_{\omega \in \mathcal{E}(p^\nu)} \rho_\nu(\omega)^4 \\ &\ll X^{\lambda+\iota} p^{-(\lambda+\iota)B/2^{2n+2}} \\ &= X^{\lambda-\iota-\iota/2^{2n+2}} \\ &= o(X^{\lambda-\iota}). \end{aligned}$$

By our choice of $\mathfrak{a} \in \mathbb{D}$, and the discussions at the beginning of Section 4.2, there is consequently some value of h with $1 \le h \le \nu$ with the property that

$$\mathfrak{J}(X) \ll \nu \, |\mathcal{E}(p)|^{2s-3+h} K_{h,h}^h.$$

By Hölder's inequality, we have

$$K_{h,h}^h \le |\mathcal{E}(p^u)|^{4t-1} K_{h,h+u}^h,$$

and consequently

$$\mathfrak{J}(X) \ll \nu \, |\mathcal{E}(p)|^{6t-3+h+u(4t-1)} K_{h,h+u}^h. \tag{4.3.1}$$

We define a sequence of indices by the following recurrence relations:

$$a_0 = h, \quad b_0 = h+u, \quad a_m = b_{m-1}, \quad b_m = 2b_{m-1} - h + 1.$$

For convenience we note that $b_m = 2^m(u+1) + h - 1$. By Lemma 4.2.4, while

$p^{b_m} < X$ we have

$$\begin{aligned} K^h_{a_m,b_m} &\ll p^{\delta(2b_m-a_m-h+1)}(X/p^{b_m})^{(\lambda+\iota)/2}(K^h_{a_{m+1},b_{m+1}})^{1/2} \\ &\ll p^{3\cdot 2^{m-1}(u+1)\delta}(X/p^{b_m})^{(\lambda+\iota)/2}(K^h_{a_{m+1},b_{m+1}})^{1/2}, \end{aligned}$$

for $m \geq 1$, and by iterating this relation and using the definition of n, we conclude that

$$\begin{aligned} K^h_{h,h+u} &\ll p^{\delta(2u+1+3(u+1)(n-1)/2)-n(\lambda+\iota)(u+1)/2}X^{(\lambda+\iota)(1-1/2^n)}(K^h_{a_n,b_n})^{1/2^n} \\ &\ll p^{\delta u/2-6t(u+1)}X^{(\lambda+\iota)(1-1/2^n)}(K^h_{a_n,b_n})^{1/2^n}. \end{aligned}$$

Substituting this into (4.3.1), we see that

$$\begin{aligned} \mathfrak{J}(X) &\ll \nu\,|\mathcal{E}(p)|^{6t-3+h+u(4t-1)}\,p^{\delta u/2-6t(u+1)}X^{(\lambda+\iota)(1-1/2^n)}(K^h_{a_n,b_n})^{1/2^n} \\ &\ll p^{\nu-u(2t+1)+\delta u/2}X^{(\lambda+\iota)(1-1/2^n)}(K^h_{a_n,b_n})^{1/2^n}\log X. \end{aligned} \tag{4.3.2}$$

By (4.2.1), we may apply the upper bound from Lemma 4.2.2, which tells us that

$$\begin{aligned} K^h_{a_n,b_n} &\ll (X/p^{a_n/3+2b_n/3})^{\lambda+\iota} \\ &= (X/p^{h-1+5\cdot 2^{n-1}(u+1)/3})^{\lambda+\iota}, \end{aligned}$$

and hence

$$(K^h_{a_n,b_n})^{1/2^n} \ll X^{(\lambda+\iota)/2^n}p^{-5u\lambda/6}.$$

Combining this with (4.3.2), and using (4.2.2) we obtain

$$\begin{aligned} \mathfrak{J}(X) &\ll p^{\nu-2tu-u+\delta u/2-5u(\Lambda+3\delta)/6}X^{\lambda+\iota+\epsilon} \\ &\ll p^{-2\iota B-u}X^{\lambda+\iota+\epsilon} \\ &\ll X^{\lambda-\iota-1/2^{n+2}+\epsilon} = o(X^{\lambda-\iota}), \end{aligned}$$

which provides the required contradiction and completes the proof of Theorem 4.1.1.

Chapter 5

Efficient Congruencing with Ellipsephic Sets: the general case

The work in this chapter is based on the author's upcoming paper [7].

5.1 Introduction

In this chapter, we extend the results of Chapter 4 to the case of general degree k. Much of the notation used here is defined in Section 4.1, although some generalisations are required. Consider polynomials $\phi_1, \dots, \phi_k \in \mathbb{Z}[z]$ which resemble those in the Vinogradov system in the following way: for $c \in \mathbb{N}$, we say that the system $\boldsymbol{\phi} = (\phi_1, \dots, \phi_k)$ is *p^c-spaced* if

$$\phi_j(z) \equiv z^j \pmod{p^c} \quad (1 \leq j \leq k).$$

For a system $\boldsymbol{\phi} \in \mathbb{Z}[z]^k$ of p^c-spaced polynomials , and a sequence $\mathfrak{a} = (\mathfrak{a}_n)_{n \in \mathcal{E}}$ of complex weights, we let

$$J_{s,k}(X) = J_{s,k}(X; \mathfrak{a}, \boldsymbol{\phi}) = \oint \Big| \sum_{x \in \mathcal{E}(X)} \mathfrak{a}_x e\big(\alpha_1\phi_1(x) + \ldots + \alpha_k\phi_k(x)\big)\Big|^{2s} d\boldsymbol{\alpha},$$

and observe that $J_{s,k}(X)$ counts the solutions, in positive integers $x_i, y_i \in \mathcal{E}(X)$, to the system

$$\phi_j(x_1) + \ldots + \phi_j(x_s) = \phi(y_1) + \ldots + \phi(y_s), \quad (1 \le j \le k),$$

where each solution is counted with weight $\mathfrak{a}_{\boldsymbol{x}}\overline{\mathfrak{a}_{\boldsymbol{y}}} = \mathfrak{a}_{x_1} \ldots \mathfrak{a}_{x_s}\overline{\mathfrak{a}_{y_1} \ldots \mathfrak{a}_{y_s}}$.

The main theorem of this chapter, of which a special case was provided in the form of Theorem 1.3.2, provides the following upper bound for $J_{s,k}(X)$.

Theorem 5.1.1. *For natural numbers k and t with $t \ge 2$, and for p a suitably large prime, let $\mathcal{E}$ be a $(p,t)^*$-ellipsephic set, and write $Y = \#\mathcal{E}(X)$. Let $\boldsymbol{\phi} \in \mathbb{Z}[z]^k$ be a system of p^c-spaced polynomials for some suitably large c. Then for $s \ge tk(k+1)/2$, we have*

$$J_{s,k}(X) \ll Y^{s-tk(k+1)/2}X^{\epsilon}\bigg(\sum_{x\in\mathcal{E}(X)} |\mathfrak{a}_x|^2\bigg)^s.$$

Corollary 5.1.2. *For natural numbers k and t with $t \ge 2$, and for p a suitably large prime, let $\mathcal{E}$ be a $(p,t)^*$-ellipsephic set, and write $Y = \#\mathcal{E}(X)$. For $s \ge tk(k+1)/2$, the number of solutions to (1.1.5) with $x_i, y_i \in \mathcal{E}(X)$ for all i is $O(Y^{2s-tk(k+1)/2}X^{\epsilon})$.*

Proof. This is the case where $\phi_j(z) = z^j$ for $1 \le j \le k$, and $\mathfrak{a}_x = 1$ for all $x \in \mathcal{E}$. □

At our critical value of $s = tk(k+1)/2$, we have

$$J_{s,k}(X) \ll X^{\epsilon}\bigg(\sum_{x\in\mathcal{E}(X)} |\mathfrak{a}_x|^2\bigg)^s,$$

whereas if we take $\mathfrak{a}_x = 0$ for $x \notin \mathcal{E}$ in the classical weighted version of Vinogradov's mean value theorem at $s = tk(k+1)/2$, we obtain

$$J_{s,k}(X) \ll Y^{(t-1)k(k+1)/2}X^{\epsilon}\bigg(\sum_{x\in\mathcal{E}(X)} |\mathfrak{a}_x|^2\bigg)^s,$$

so we see that, as in the quadratic case, we have achieved a power saving in Y by using the additive structure of our ellipsephic sets, rather than simply their density.

An important area for future consideration is the application of the results of this chapter to Waring's problem, in which we seek to find $s = s(k)$ such that any $n \in \mathbb{N}$ may be written in the form (1.1.1) with $x_1, \dots, x_s \in \mathcal{E}$. As in Section 4.1, we are able to prove a lower bound for $N_{s,k}(X) = N_{s,k}^{\mathcal{E}}(X)$, defined as the number of positive integers up to X which have a representation in such a form.

Corollary 5.1.3. *For natural numbers k and t with $t \geq 2$, and for p a suitably large prime, let $\mathcal{E}$ be a $(p,t)^*$-ellipsephic set. Assume that $Y = \#\mathcal{E}(X) \gg X^{1/t}$. Then for $s \geq tk(k+1)/2$ we have*

$$N_{s,k}(X) \gg X^{1-\epsilon}.$$

Proof. As in Corollary 4.1.4, we apply Cauchy's inequality and Corollary 5.1.2 to obtain the bound

$$N_{s,k}(X) \gg Y^{t(k+1)/2} X^{(1-k)/2-\epsilon},$$

and then use our assumption on the size of Y to deduce that

$$N_{s,k}(X) \gg X^{1-\epsilon},$$

as required. □

The proof of Theorem 5.1.1 uses Wooley's nested efficient congruencing method and closely follows the argument of [71], with suitable adjustments for our ellipsephic situation. In Section 5.2, we provide preliminary notation and formulate an alternative theorem (Theorem 5.2.1), which we prove by induction in the next four sections. Specifically, in Section 5.3, which is the main point of divergence from the work of Wooley, we use the additive properties of our $(p,t)^*$-ellipsephic sets to prove the base case ($k = 1$) of Theorem 5.2.1, using a "lifting" argument similar to that in Section 4.2. In Section 5.4 we introduce a "hierarchy" of small constants to support the rest of the chapter, and prove some basic results, and in Section 5.5 we use the inductive hypothesis to prove a series of lemmata which form the backbone of our iteration. In Section 5.6 we complete the proof of Theorem 5.2.1, hypothesising that a certain quantity is strictly greater than zero and deriving a contradiction. Finally, in Section 5.7 we use Theorem 5.2.1 to deduce Theorem 5.1.1.

5.2 Preliminaries

In a similar way to that used in Chapter 4, we normalise our exponential sums as follows. For a sequence $\mathfrak{a} = (\mathfrak{a}_n)_{n\in\mathcal{E}}$ of complex weights satisfying $0 < \sum_{n\in\mathcal{E}} |\mathfrak{a}_n| < \infty$, we let

$$\rho_0 = \bigg(\sum_{x\in\mathcal{E}} |\mathfrak{a}_x|^2 \bigg)^{1/2},$$

and for $\boldsymbol{\alpha} \in [0,1]^k$, we let

$$f(\boldsymbol{\alpha}) = f(\boldsymbol{\alpha}; \mathfrak{a}) = \rho_0^{-1} \sum_{x\in\mathcal{E}} \mathfrak{a}_x e(\psi(x;\boldsymbol{\alpha})),$$

where $\psi(x;\boldsymbol{\alpha}) = \alpha_1\phi_1(x) + \ldots + \alpha_k\phi_k(x)$. A bound of the form

$$J_{s,k}(X) \ll X^{\Delta} \bigg(\sum_{x\in\mathcal{E}} |\mathfrak{a}_x|^2 \bigg)^s$$

therefore follows directly from one of the form

$$\oint |f(\boldsymbol{\alpha})|^{2s}\, d\boldsymbol{\alpha} \ll X^{\Delta}.$$

We also wish to define the restriction of $f(\boldsymbol{\alpha})$ to congruence classes modulo various powers of p. For $a \in \mathbb{N}$ and $\xi \in \mathcal{E}(p^a)$, let

$$\rho_a(\xi) = \bigg(\sum_{\substack{x\in\mathcal{E} \\ x\equiv\xi \pmod{p^a}}} |\mathfrak{a}_x|^2 \bigg)^{1/2}$$

and

$$f_a(\boldsymbol{\alpha}, \xi) = \rho_a(\xi)^{-1} \sum_{\substack{x\in\mathcal{E} \\ x\equiv\xi \pmod{p^a}}} \mathfrak{a}_x e(\psi(x;\boldsymbol{\alpha})). \tag{5.2.1}$$

As in Chapter 4, under this normalisation we may assume that every $\mathfrak{a}_x$ is real, non-negative and at most one; we write

$$\mathbb{D} = \bigg\{ \mathfrak{a} \;\bigg|\; 0 \le \mathfrak{a}_n \le 1 \text{ for all } n \in \mathcal{E} \text{ and } 0 < \sum_{n\in\mathcal{E}} \mathfrak{a}_n < \infty \bigg\},$$

and we work with $\mathfrak{a} \in \mathbb{D}$.

For later convenience, for any ξ we interpret $\rho_0(\xi)$ to be ρ_0 and $f_0(\boldsymbol{\alpha}, \xi)$ to be $f(\boldsymbol{\alpha})$, and we observe that for $a \in \mathbb{N}$, we have

$$\sum_{\xi \in \mathcal{E}(p^a)} \rho_a(\xi)^2 = \rho_0^2,$$

and for $a, b \in \mathbb{N}$ with $a \leq b$,

$$\sum_{\substack{\xi' \in \mathcal{E}(p^b) \\ \xi' \equiv \xi \pmod{p^a}}} \rho_b(\xi')^2 = \rho_a(\xi)^2.$$

Our strategy for counting solutions to the system of equations we are interested in involves studying congruences modulo suitably large powers of p, and as such we recall the notation

$$\oint_{p^B} F(\boldsymbol{\alpha})\, d\boldsymbol{\alpha} = p^{-kB} \sum_{1 \leq u_1 \leq p^B} \dots \sum_{1 \leq u_k \leq p^B} F(\boldsymbol{u}/p^B),$$

and define

$$U_{s,k}^B(\mathfrak{a}) = \oint_{p^B} |f(\boldsymbol{\alpha})|^{2s}\, d\boldsymbol{\alpha},$$

which counts solutions to the system of congruences

$$\sum_{i=1}^{s} \big(\phi_j(x_i) - \phi_j(y_i)\big) \equiv 0 \pmod{p^B}, \quad (1 \leq j \leq k) \tag{5.2.2}$$

with $\boldsymbol{x}, \boldsymbol{y} \in \mathcal{E}^s$, where each solution is counted with weight $\rho_0^{-2s} \mathfrak{a}_{\boldsymbol{x}} \mathfrak{a}_{\boldsymbol{y}}$. We also wish to count solutions to (5.2.2) with further congruence restrictions on our variables, so for $H \in \mathbb{N}$, we let

$$U_{s,k}^{B,H}(\mathfrak{a}) = \rho_0^{-2} \sum_{\xi \in \mathcal{E}(p^H)} \rho_H(\xi)^2 \oint_{p^B} |f_H(\boldsymbol{\alpha}, \xi)|^{2s}\, d\boldsymbol{\alpha}.$$

The integral on the right-hand side imposes the additional condition that $\boldsymbol{x} \equiv \boldsymbol{y} \equiv \xi \pmod{p^H}$, and the solutions are now counted with weight $\rho_H(\xi)^{-2s} \mathfrak{a}_{\boldsymbol{x}} \mathfrak{a}_{\boldsymbol{y}}$.

We observe that, for $H \in \mathbb{N}$, we have

$$f(\boldsymbol{\alpha}) = \rho_0^{-1} \sum_{\xi \in \mathcal{E}(p^H)} \rho_H(\xi) f_H(\boldsymbol{\alpha}, \xi),$$

so, by Hölder's inequality,

$$\begin{aligned} |f(\boldsymbol{\alpha})|^{2s} &\leq \rho_0^{-2s} \Big(\sum_{\xi \in \mathcal{E}(p^H)} 1 \Big)^s \Big(\sum_{\xi \in \mathcal{E}(p^H)} \rho_H(\xi)^2 \Big)^{s-1} \sum_{\xi \in \mathcal{E}(p^H)} \rho_H(\xi)^2 \left| f_H(\boldsymbol{\alpha}, \xi) \right|^{2s} \\ &\ll \rho_0^{-2} q^{sH} \sum_{\xi \in \mathcal{E}(p^H)} \rho_H(\xi)^2 \left| f_H(\boldsymbol{\alpha}, \xi) \right|^{2s}, \end{aligned}$$

where we have written $q = \#\mathcal{E}(p)$. Consequently, we have

$$U_{s,k}^B(\mathfrak{a}) \ll q^{sH} U_{s,k}^{B,H}(\mathfrak{a}). \tag{5.2.3}$$

We may now ask for the minimal value of λ such that

$$U_{s,k}^B(\mathfrak{a}) \ll (q^H)^{\lambda+\epsilon} U_{s,k}^{B,H}(\mathfrak{a})$$

as uniformly as possible in the parameters $(\mathfrak{a}, \boldsymbol{\phi}, B)$, and observe that (5.2.3) implies that $\lambda \leq s$.

For $\tau > 0$, let $\Phi_\tau(B)$ denote the set of systems $\boldsymbol{\phi}$ which are p^c-spaced for some $c \geq \tau B$. We deduce from (5.2.3) that for $\boldsymbol{\phi} \in \Phi_\tau(B)$, we have

$$\sup_{\mathfrak{a} \in \mathbb{D}} \frac{\log \left(U_{s,k}^B(\mathfrak{a}) / U_{s,k}^{B,H}(\mathfrak{a}) \right)}{\log q^H} \leq s$$

for all $H \in \mathbb{N}$. Now consider the particular choice of $\mathfrak{b} \in \mathbb{D}$ with $\mathfrak{b} = 0$ whenever $n \not\equiv 0 \pmod{p^H}$. We have $U_{s,k}^B(\mathfrak{b}) = U_{s,k}^{B,H}(\mathfrak{b})$, and consequently

$$\sup_{\mathfrak{a} \in \mathbb{D}} \frac{\log \left(U_{s,k}^B(\mathfrak{a}) / U_{s,k}^{B,H}(\mathfrak{a}) \right)}{\log q^H} \geq 0.$$

Given $s, k \in \mathbb{N}$ and $\tau > 0$, we let $H = \lceil B/k \rceil$ and let

$$\lambda^*(s, k; \tau) = \limsup_{B \to \infty} \sup_{\boldsymbol{\phi} \in \Phi_\tau(B)} \sup_{\mathfrak{a} \in \mathbb{D}} \frac{\log \left(U_{s,k}^B(\mathfrak{a}) / U_{s,k}^{B,H}(\mathfrak{a}) \right)}{\log q^H},$$

and

$$\lambda(s,k) = \limsup_{\tau\to 0} \lambda^*(s,k;\tau). \tag{5.2.4}$$

We then have $0 \le \lambda^*(s,k;\tau) \le s$ and consequently $0 \le \lambda(s,k) \le s$. This leads us to the statement of a key result to be used in the proof of Theorem 5.1.1.

Theorem 5.2.1. *For natural numbers k and t with $t \ge 2$, and for p a suitably large prime, let $\mathcal{E}$ be a $(p,t)^*$-ellipsephic set. Then $\lambda(tk(k+1)/2,k) = 0$.*

As a corollary, we obtain

Corollary 5.2.2. *For natural numbers k and t with $t \ge 2$, and for p a suitably large prime, let $\mathcal{E}$ be a $(p,t)^*$-ellipsephic set. Let $\tau > 0$ and $\epsilon > 0$, and let B be sufficiently large in terms of k, τ and ϵ. Set $s = tk(k+1)/2$ and $H = \lceil B/k \rceil$. Then for all $\boldsymbol{\phi} \in \Phi_\tau(B)$ and $\mathfrak{a} \in \mathbb{D}$, we have*

$$U_{s,k}^B(\mathfrak{a}) \ll q^{H\epsilon} U_{s,k}^{B,H}(\mathfrak{a}).$$

Proof. By the definition of $\lambda^*(s,k;\tau)$, we have, for sufficiently large B, the bound

$$U_{s,k}^B(\mathfrak{a}) \ll (q^H)^{\lambda^*(s,k;\tau)+\epsilon} U_{s,k}^{B,H}(\mathfrak{a}).$$

Allowing τ to tend to zero and applying Theorem 5.2.1 gives the result. □

We introduce some final definitions. For $a,b,c,\nu \in \mathbb{N}$, and for $0 \le r \le k$ and $R = tr(r+1)/2$, we let

$$K_{a,b,c}^{r,\boldsymbol{\phi}}(\mathfrak{a};\xi,\eta) = \oint_{p^B} \left| f_a(\boldsymbol{\alpha},\xi)^{2R} f_b(\boldsymbol{\alpha},\eta)^{2s-2R} \right| d\boldsymbol{\alpha}$$

and

$$K_{a,b,c}^{r,\boldsymbol{\phi},\nu}(\mathfrak{a}) = \rho_0^{-4} \sum_{\xi\in\mathcal{E}(p^a)} \sum_{\substack{\eta\in\mathcal{E}(p^b)\\ \xi\not\equiv\eta \pmod{p^\nu}}} \rho_a(\xi)^2 \rho_b(\eta)^2 K_{a,b,c}^{r,\boldsymbol{\phi}}(\mathfrak{a};\xi,\eta). \tag{5.2.5}$$

Note that $K_{a,b,c}^{r,\boldsymbol{\phi}}(\mathfrak{a};\xi,\eta)$ counts solutions $(\boldsymbol{x},\boldsymbol{y},\boldsymbol{u},\boldsymbol{v}) \in \mathcal{E}^{2s}$ to the congruences

$$\sum_{i=1}^{R} \big(\phi_j(x_i) - \phi_j(y_i)\big) \equiv \sum_{l=1}^{s-R} \big(\phi_j(u_l) - \phi_j(v_l)\big) \pmod{p^B} \quad (1 \le j \le k),$$

with $\boldsymbol{x} \equiv \boldsymbol{y} \equiv \xi \pmod{p^a}$ and $\boldsymbol{u} \equiv \boldsymbol{v} \equiv \eta \pmod{p^b}$, where each solution is counted with weight $\rho_a(\xi)^{-2R}\rho_b(\eta)^{2R-2s}\mathfrak{a}_{\boldsymbol{x}}\mathfrak{a}_{\boldsymbol{y}}\mathfrak{a}_{\boldsymbol{u}}\mathfrak{a}_{\boldsymbol{v}}$.

We are also interested in normalised versions of these mean values, so for $\Delta \geq 0$ we define

$$\widetilde{K}^{r,\boldsymbol{\phi},\nu}_{a,b,c}(\mathfrak{a})_\Delta = \left(\frac{K^{r,\boldsymbol{\phi},\nu}_{a,b,c}(\mathfrak{a})}{q^{\Delta H}U^{B,H}_{s,k}(\mathfrak{a})}\right)^{\frac{k-1}{r(k-r)}}. \tag{5.2.6}$$

We now prove some auxiliary results giving bounds on the above-defined mean values.

Lemma 5.2.3. *For $s, k \in \mathbb{N}$ and $p > k$ a prime, let $\mathcal{E}$ be a $(p,t)^*$-ellipsephic set. Let $0 < \epsilon < \tau < \delta < 1$, and let B be sufficiently large in terms of s, k and ϵ. Set $H = \lceil B/k \rceil$. Then for all $\boldsymbol{\phi} \in \Phi_\tau(B)$, for all $\mathfrak{a} \in \mathbb{D}$, and for all $h \in \mathbb{N}_0$ with $h \leq (1-\delta)H$, we have*

$$U^{B,h}_{s,k}(\mathfrak{a}) \ll (q^{H-h})^{\lambda(s,k)+\epsilon}U^{B,H}_{s,k}(\mathfrak{a}).$$

Proof. The integral within the definition of $U^{B,h}_{s,k}(\mathfrak{a})$ counts solutions to the system of congruences (5.2.2) with $\boldsymbol{x}, \boldsymbol{y} \in \mathcal{E}^s$ and $\boldsymbol{x} \equiv \boldsymbol{y} \equiv \xi \pmod{p^h}$, with weights $\rho_H(\xi)^{-2s}\mathfrak{a}_{\boldsymbol{x}}\mathfrak{a}_{\boldsymbol{y}}$. As in [71, Lemma 4.1], we make use of some linear algebra to transform this situation into one in which we have a set of p^{c+h}-spaced polynomials

$$\Phi_j(z) = z^j + p^{c+h}z^{k+1}\Upsilon_j(z) \quad (1 \leq j \leq k),$$

for some $\Upsilon_j \in \mathbb{Z}[z]$, satisfying

$$\sum_{i=1}^{s}\Phi_j(x_i) \equiv \sum_{i=1}^{s}\Phi_j(y_i) \pmod{p^{B-kh}} \quad (1 \leq j \leq k)$$

whenever $\boldsymbol{x}, \boldsymbol{y}$ forms a solution to the original system of congruences counted by $U^{B,h}_{s,k}(\mathfrak{a})$.

The fact that $h \leq (1-\delta)H$ allows us to assume that $B - kh$ is sufficiently large with respect to s, k and ϵ, and consequently the definition (5.2.4) yields

$$U^{B-kh}_{s,k}(\mathfrak{c}) \ll (q^{H-h})^{\lambda(s,k)+\epsilon}U^{B-kh,H-h}_{s,k}(\mathfrak{c}),$$

where $\mathfrak{c}$ is an auxiliary set of weights defined by $\mathfrak{c}_u = \mathfrak{a}_{p^hu+\xi}\, e(\psi(p^hu+\xi;\boldsymbol{\alpha}))$.

Rearranging, and using orthogonality, we obtain the conclusion. □

Lemma 5.2.4. *For $s, k \in \mathbb{N}$ and $p > k$ a prime, let $\mathcal{E}$ be a $(p,t)^*$-ellipsephic set. Let $0 < \epsilon < \tau < \delta < 1$, and let B be sufficiently large in terms of s, k and ϵ. Set $H = \lceil B/k \rceil$ and let $r, \nu \in \mathbb{N}_0$ with $1 \le r \le k-1$. Suppose that $0 < \Lambda \le \lambda(s,k)$. Then for all $\boldsymbol{\phi} \in \Phi_\tau(B)$, for all $\mathfrak{a} \in \mathbb{D}$, and for all $a, b \in \mathbb{N}_0$ with $\max\{a,b\} \le (1-\delta)H$, we have*

$$\widetilde{K}^{r,\boldsymbol{\phi},\nu}_{a,b,c}(\mathfrak{a})_\Lambda \ll (q^H)^{\lambda(s,k)-\Lambda+\epsilon}.$$

Proof. As in [71, Lemma 4.2], we apply Hölder's inequality to see that

$$\rho_0^{-4} \sum_{\xi \in \mathcal{E}(p^a)} \sum_{\eta \in \mathcal{E}(p^b)} \rho_a(\xi)^2 \rho_b(\eta)^2 \oint_{p^B} \left| f_a(\boldsymbol{\alpha}, \xi)^{2R} f_b(\boldsymbol{\alpha}, \eta)^{2s-2R} \right| d\boldsymbol{\alpha} \le I_1^{R/s} I_2^{1-R/s},$$

where

$$I_1 = \rho_0^{-4} \sum_{\xi \in \mathcal{E}(p^a)} \sum_{\eta \in \mathcal{E}(p^b)} \rho_a(\xi)^2 \rho_b(\eta)^2 \oint_{p^B} |f_a(\boldsymbol{\alpha}, \xi)|^{2s} \, d\boldsymbol{\alpha} = U^{B,a}_{s,k}(\mathfrak{a}),$$

and

$$I_2 = \rho_0^{-4} \sum_{\xi \in \mathcal{E}(p^a)} \sum_{\eta \in \mathcal{E}(p^b)} \rho_a(\xi)^2 \rho_b(\eta)^2 \oint_{p^B} |f_b(\boldsymbol{\alpha}, \eta)|^{2s} \, d\boldsymbol{\alpha} = U^{B,b}_{s,k}(\mathfrak{a}).$$

Applying Lemma 5.2.3, we see that

$$K^{r,\boldsymbol{\phi},\nu}_{a,b,c}(\mathfrak{a}) \le \left((q^{H-a})^{R/s} (q^{H-b})^{1-R/s} \right)^{\lambda(s,k)+\epsilon} U^{B,H}_{s,k}(\mathfrak{a}),$$

and therefore, using the definition (5.2.6), that

$$\begin{aligned} \left(\widetilde{K}^{r,\boldsymbol{\phi},\nu}_{a,b,c}(\mathfrak{a})_\Lambda \right)^{r(k-r)/(k-1)} &\le \left((q^{H-a})^{R/s} (q^{H-b})^{1-R/s} \right)^{\lambda(s,k)+\epsilon} q^{-\Lambda H} \\ &\le q^{H(\lambda(s,k)-\Lambda+\epsilon)}. \end{aligned}$$

Since $1 \le r \le k-1$, we have $r(k-r) \ge k-1$, and consequently

$$\widetilde{K}^{r,\boldsymbol{\phi},\nu}_{a,b,c}(\mathfrak{a})_\Lambda \le q^{H(\lambda(s,k)-\Lambda+\epsilon)(k-1)/r(k-r)} \le q^{H(\lambda(s,k)-\Lambda+\epsilon)},$$

as claimed. □

5.3 The base case $k = 1$

In this section, we use the properties of our $(p,t)^*$-ellipsephic sets to prove that Theorem 5.2.1 holds in the case $k = 1$. The arguments resemble those used in Section 4.2, as well as the base case [71, Lemma 5.1] in the work of Wooley.

Lemma 5.3.1. *For $t \geq 2$ an integer, and p a sufficiently large prime, let $\mathcal{E}$ be a $(p,t)^*$-ellipsephic set. Then $\lambda(t,1) = 0$.*

Proof. Let $0 < \tau < 1$, and let $B \in \mathbb{N}$ be sufficiently large in terms of τ. Fix any $\mathfrak{a} \in \mathbb{D}$ and any $\phi \in \Phi_\tau(B)$, so that by definition we have $\phi(z) = z + p^c\psi(z)$ for some $c \geq \tau B$ and some $\psi \in \mathbb{Z}[z]$. Then $U_{t,1}^B(\mathfrak{a})$ counts solutions to the congruence

$$\sum_{i=1}^{t} \big(\phi(x_i) - \phi(y_i)\big) \equiv 0 \ (\mathrm{mod}\ p^B) \tag{5.3.1}$$

with $\boldsymbol{x}, \boldsymbol{y} \in \mathcal{E}^t$, and where each solution is counted with weight $\rho_0^{-2t}\mathfrak{a}_{\boldsymbol{x}}\mathfrak{a}_{\boldsymbol{y}}$. We may rewrite (5.3.1) in the form

$$\sum_{i=1}^{t} \big(x_i + p^c\psi(x_i)\big) \equiv \sum_{i=1}^{t} \big(y_i + p^c\psi(y_i)\big) \ (\mathrm{mod}\ p^B), \tag{5.3.2}$$

allowing us to deduce that

$$\sum_{i=1}^{t} x_i \equiv \sum_{i=1}^{t} y_i \ (\mathrm{mod}\ p^{c_1}), \tag{5.3.3}$$

where we write $c_1 = \min\{B, c\}$. This is effectively a 'free' condition which was already contained in our original congruence (5.3.1). For $d \in \mathbb{N}$ and weights $\mathfrak{b} \in \mathbb{D}$ we define

$$G_d(\mathfrak{b}) = \oint_{p^d} \Big| \sum_{\boldsymbol{x} \in \mathcal{E}^t} \mathfrak{b}_{\boldsymbol{x}} e\big(\beta(x_1 + \ldots + x_t)\big) \Big|^2 d\beta.$$

We fix the weights appearing in this definition to be

$$\mathfrak{b}_x = \rho_0^{-1}\mathfrak{a}_x e(\alpha\phi(x)).$$

Then $G_{c_1}(\mathfrak{b})$ encodes the number of solutions to (5.3.3), counted with weights $\rho_0^{-2t}\mathfrak{a}_{\boldsymbol{x}}\mathfrak{a}_{\boldsymbol{y}} e\Big(\alpha \sum_{i=1}^{t} \big(\phi(x_i) - \phi(y_i)\big)\Big)$, and consequently we may insert the con-

dition (5.3.3) into our original congruence in the form

$$U_{t,1}^B(\mathfrak{a}) = \oint_{p^B} \Big| \rho_0^{-1} \sum_{x \in \mathcal{E}} \mathfrak{a}_x e(\alpha\phi(x)) \Big|^{2t} d\alpha = \oint_{p^B} G_{c_1}(\mathfrak{b})\, d\alpha.$$

By a slight adaptation of Lemma 4.2.3, we have

$$G_{c_1}(\mathfrak{b}) \ll p^\epsilon \sum_{\boldsymbol{u} \in \mathcal{E}(p^{c_1})^t} \Big| \sum_{\substack{\boldsymbol{x} \in \mathcal{E}^t \\ \boldsymbol{x} \equiv \boldsymbol{u} \pmod{p^{c_1}}}} \mathfrak{b}_{\mathbf{x}} \Big|^2,$$

for any $\epsilon > 0$, and consequently

$$U_{t,1}^B(\mathfrak{a}) \ll p^\epsilon \sum_{\boldsymbol{u} \in \mathcal{E}(p^{c_1})^t} \oint_{p^B} \Big| \sum_{\substack{\boldsymbol{x} \in \mathcal{E}^t \\ \boldsymbol{x} \equiv \boldsymbol{u} \pmod{p^{c_1}}}} \mathfrak{b}_{\mathbf{x}} \Big|^2 d\alpha,$$

where the integrand on the right-hand side now imposes the condition $\boldsymbol{x} \equiv \boldsymbol{y} \equiv \boldsymbol{u} \pmod{p^{c_1}}$. The fact that p^{c_1} divides $x_i - y_i$ implies that p^{c_1} divides $\psi(x_i) - \psi(y_i)$ for $1 \le i \le t$, and substituting this into (5.3.2) gives the congruence

$$\sum_{i=1}^t x_i \equiv \sum_{i=1}^t y_i \pmod{p^{c_2}},$$

where $c_2 = \min\{2c, B\}$. Repeating this process, we eventually reach the point at which our congruence is modulo p^{c_j} with $c_j = \min\{jc, B\} = B$, and since $c \ge \tau B$, this happens after at most $\lceil \tau^{-1} \rceil$ steps. Now

$$U_{t,1}^B(\mathfrak{a}) \ll p^\epsilon \sum_{\boldsymbol{u} \in \mathcal{E}(p^B)^t} \oint_{p^B} \Big| \sum_{\substack{\boldsymbol{x} \in \mathcal{E}^t \\ \boldsymbol{x} \equiv \boldsymbol{u} \pmod{p^B}}} \mathfrak{b}_{\mathbf{x}} \Big|^2 d\alpha,$$

so returning to the definition of the weights $\mathfrak{b}$, we obtain

$$\begin{aligned} U_{t,1}^B(\mathfrak{a}) &\ll p^\epsilon \rho_0^{-2t} \sum_{\boldsymbol{u} \in \mathcal{E}(p^B)^t} \oint_{p^B} \Big| \sum_{\substack{\boldsymbol{x} \in \mathcal{E}^t \\ \boldsymbol{x} \equiv \boldsymbol{u} \pmod{p^B}}} \mathfrak{a}_{\boldsymbol{x}} e\Big(\alpha \sum_{i=1}^t \phi(x_i)\Big) \Big|^2 d\alpha \\ &= p^\epsilon \rho_0^{-2t} \sum_{\boldsymbol{u} \in \mathcal{E}(p^B)^t} \oint_{p^B} \Big| \prod_{i=1}^t \rho_B(u_i) f_B(\alpha, u_i) \Big|^2 d\alpha. \end{aligned}$$

Using Hölder's inequality twice, we see that

$$\begin{aligned} U_{t,1}^B(\mathfrak{a}) &\ll p^\epsilon \rho_0^{-2t} \sum_{\boldsymbol{u}\in\mathcal{E}(p^B)^t} \prod_{i=1}^t \rho_B(u_i)^2 \Big(\oint_{p^B} \big|f_B(\alpha,u_i)\big|^{2t}\, d\alpha \Big)^{1/t} \\ &= p^\epsilon \rho_0^{-2t} \Bigg(\sum_{u\in\mathcal{E}(p^B)} \rho_B(u)^2 \Big(\oint_{p^B} \big|f_B(\alpha,u)\big|^{2t}\, d\alpha \Big)^{1/t} \Bigg)^t \\ &\ll p^\epsilon \rho_0^{-2} \sum_{u\in\mathcal{E}(p^B)} \rho_B(u)^2 \oint_{p^B} \big|f_B(\alpha,u)\big|^{2t}\, d\alpha = p^\epsilon U_{t,1}^{B,B}(\mathfrak{a}). \end{aligned}$$

We may assume that B is sufficiently large to give $p^\epsilon \ll q^{B\epsilon}$, and consequently we deduce that

$$\frac{\log(U_{t,1}^B(\mathfrak{a})/U_{t,1}^{B,B}(\mathfrak{a}))}{\log q^B} \ll \epsilon$$

for any $\epsilon > 0$, and hence, using the definition (5.2.4), we find that $\lambda(t,1) = 0$ as claimed. □

5.4 The hierarchy

In order to prove Theorem 5.2.1, we assume that $\Lambda = \lambda(tk(k+1)/2, k) > 0$, and work towards a contradiction. We introduce a series of small positive numbers

$$0 < \epsilon < \tau < \delta < \mu < 1, \tag{5.4.1}$$

which form a hierarchy in the sense that each element is assumed to be small enough in terms of k, Λ and the larger parameters in the inequality (5.4.1). We may then choose B large enough, in terms of $k, \Lambda, \mu, \delta, \tau$ and ϵ, to ensure that, writing $H = \lceil B/k \rceil$, we have

$$U_{s,k}^B(\mathfrak{a}) \geq (q^H)^{\Lambda-\epsilon} U_{s,k}^{B,H}(\mathfrak{a}). \tag{5.4.2}$$

By Lemma 5.2.3, we may also assume that for all $h \in \mathbb{N}_0$ with $h \leq (1-\delta)H$, and for all $\mathfrak{a}' \in \mathbb{D}$, we have

$$U_{s,k}^{B,h}(\mathfrak{a}') \leq (q^{H-h})^{\Lambda+\epsilon} U_{s,k}^{B,H}(\mathfrak{a}'). \tag{5.4.3}$$

We also fix parameters

$$\nu = \lceil 4\epsilon H \Lambda^{-1} \rceil \quad \text{and} \quad \theta = \lceil \mu H \rceil \tag{5.4.4}$$

for use in the remainder of the chapter.

Lemma 5.4.1. *We have* $U^B_{s,k}(\mathfrak{a}) \ll q^{s\nu} K^{1,\phi,\nu}_{\nu,\nu,c}(\mathfrak{a})$.

Proof. As in [71, Lemma 6.1], we have

$$\rho_0^2 f(\boldsymbol{\alpha})^2 = \sum_{\xi \in \mathcal{E}(p^\nu)} \rho_\nu(\xi) f_\nu(\boldsymbol{\alpha}, \xi)(\rho_0 f(\boldsymbol{\alpha})),$$

and, for any $\xi \in \mathcal{E}(p^\nu)$,

$$\rho_0 f(\boldsymbol{\alpha}) = \rho_\nu(\xi) f_\nu(\boldsymbol{\alpha}, \xi) + \sum_{\substack{\eta \in \mathcal{E}(p^\nu) \\ \eta \not\equiv \xi \pmod{p^\nu}}} \rho_\nu(\eta) f_\nu(\boldsymbol{\alpha}, \eta).$$

Consequently, we obtain

$$|f(\boldsymbol{\alpha})|^2 \leq T_1(\boldsymbol{\alpha}) + T_2(\boldsymbol{\alpha}),$$

where

$$T_1(\boldsymbol{\alpha}) = \rho_0^{-2} \sum_{\xi \in \mathcal{E}(p^\nu)} \rho_\nu(\xi)^2 \left| f_\nu(\boldsymbol{\alpha}, \xi) \right|^2$$

and

$$T_2(\boldsymbol{\alpha}) = \rho_0^{-2} \sum_{\xi \in \mathcal{E}(p^\nu)} \sum_{\substack{\eta \in \mathcal{E}(p^\nu) \\ \eta \not\equiv \xi \pmod{p^\nu}}} \rho_\nu(\xi) \rho_\nu(\eta) \left| f_\nu(\boldsymbol{\alpha}, \xi) f_\nu(\boldsymbol{\alpha}, \eta) \right|.$$

We now apply Hölder's inequality to obtain

$$T_1(\boldsymbol{\alpha})^s \leq \rho_0^{-2} \sum_{\xi \in \mathcal{E}(p^\nu)} \rho_\nu(\xi)^2 \left| f_\nu(\boldsymbol{\alpha}, \xi) \right|^{2s}$$

and

$$T_2(\boldsymbol{\alpha})^s \leq \rho_0^{-4} q^{s\nu} \sum_{\xi \in \mathcal{E}(p^\nu)} \sum_{\substack{\eta \in \mathcal{E}(p^\nu) \\ \eta \not\equiv \xi \pmod{p^\nu}}} \rho_\nu(\xi)^2 \rho_\nu(\eta)^2 \left| f_\nu(\boldsymbol{\alpha}, \xi)^{2t} f_\nu(\boldsymbol{\alpha}, \eta)^{2s-2t} \right|,$$

so that

$$\begin{aligned} U_{s,k}^B(\mathfrak{a}) = \oint_{p^B} |f(\boldsymbol{\alpha})|^{2s} \, d\boldsymbol{\alpha} &\ll \oint_{p^B} T_1(\boldsymbol{\alpha})^s \, d\boldsymbol{\alpha} + \oint_{p^B} T_2(\boldsymbol{\alpha})^s \, d\boldsymbol{\alpha} \\ &\ll U_{s,k}^{B,\nu}(\mathfrak{a}) + q^{s\nu} K_{\nu,\nu,c}^{1,\boldsymbol{\phi},\nu}(\mathfrak{a}). \end{aligned}$$

By Lemma 5.2.3, and using the definition of ν, we have

$$\begin{aligned} U_{s,k}^{B,\nu}(\mathfrak{a}) &\ll (q^{H-\nu})^{\Lambda+\epsilon} U_{s,k}^{B,H}(\mathfrak{a}) \\ &\ll q^{-2\epsilon H} (q^H)^{\Lambda-\epsilon} U_{s,k}^{B,H}(\mathfrak{a}), \end{aligned}$$

and by (5.4.2), this implies

$$U_{s,k}^{B,\nu}(\mathfrak{a}) \ll q^{-2\epsilon H} U_{s,k}^B(\mathfrak{a}),$$

so that

$$U_{s,k}^B(\mathfrak{a}) \ll q^{s\nu} K_{\nu,\nu,c}^{1,\boldsymbol{\phi},\nu}(\mathfrak{a})$$

as claimed. □

Lemma 5.4.2. *For $a, b \in \mathbb{N}_0$ with $a \leq b$, and $w > 0$ and $\xi \in \mathcal{E}$, we have*

$$\rho_a(\xi)^2 \left| f_a(\boldsymbol{\alpha}, \xi) \right|^{2w} \leq q^{w(b-a)} \sum_{\substack{\zeta \in \mathcal{E}(p^b) \\ \zeta \equiv \xi \ (mod\ p^a)}} \rho_b(\zeta)^2 \left| f_b(\boldsymbol{\alpha}, \zeta) \right|^{2w}.$$

Proof. Apply Hölder's inequality, as in [71, Lemma 6.2], to see that

$$\begin{aligned} \rho_a(\xi)^{2w} \left| f_a(\boldsymbol{\alpha}, \xi) \right|^{2w} &= \left| \sum_{\substack{\zeta \in \mathcal{E}(p^b) \\ \zeta \equiv \xi \pmod{p^a}}} \rho_b(\zeta) f_b(\boldsymbol{\alpha}, \zeta) \right|^{2w} \\ &\leq U_1^w U_2^{w-1} \sum_{\substack{\zeta \in \mathcal{E}(p^b) \\ \zeta \equiv \xi \pmod{p^a}}} \rho_b(\zeta)^2 \left| f_b(\boldsymbol{\alpha}, \zeta) \right|^{2w}, \end{aligned}$$

where

$$U_1 = \sum_{\substack{\zeta \in \mathcal{E}(p^b) \\ \zeta \equiv \xi \pmod{p^a}}} 1 \quad = q^{b-a},$$

and

$$U_2 = \sum_{\substack{\zeta \in \mathcal{E}(p^b) \\ \zeta \equiv \xi \pmod{p^a}}} \rho_b(\zeta)^2 = \rho_a(\xi)^2.$$

We therefore conclude that

$$\rho_a(\xi)^2 \left| f_a(\boldsymbol{\alpha}, \xi) \right|^{2w} \leq q^{w(b-a)} \sum_{\substack{\zeta \in \mathcal{E}(p^b) \\ \zeta \equiv \xi \pmod{p^a}}} \rho_b(\zeta)^2 \left| f_b(\boldsymbol{\alpha}, \zeta) \right|^{2w},$$

as required. □

Lemma 5.4.3. *We have* $U_{s,k}^B(\mathfrak{a}) \ll q^{s\theta} K_{\theta,\theta,c}^{1,\boldsymbol{\phi},\nu}(\mathfrak{a})$.

Proof. By Lemma 5.4.1, we have

$$U_{s,k}^B(\mathfrak{a}) \ll q^{s\nu} K_{\nu,\nu,c}^{1,\boldsymbol{\phi},\nu}(\mathfrak{a}). \tag{5.4.5}$$

As in [71, Lemma 6.3], we apply Lemma 5.4.2 twice to obtain

$$\begin{aligned} K_{\nu,\nu,c}^{1,\boldsymbol{\phi}}(\mathfrak{a}; \xi, \eta) &= \oint_{p^B} \left| f_\nu(\boldsymbol{\alpha}, \xi)^{2t} f_\nu(\boldsymbol{\alpha}, \eta)^{2s-2t} \right| d\boldsymbol{\alpha} \\ &\leq \rho_\nu(\xi)^{-2} \rho_\nu(\eta)^{-2} q^{s(\theta-\nu)} \sum_{\substack{\xi', \eta' \in \mathcal{E}(p^\theta) \\ (\xi', \eta') \equiv (\xi, \eta) \pmod{p^\nu}}} \rho_\theta(\xi')^2 \rho_\theta(\eta')^2 K_{\theta,\theta,c}^{1,\boldsymbol{\phi}}(\mathfrak{a}; \xi', \eta'). \end{aligned}$$

Consequently,

$$K_{\nu,\nu,c}^{1,\boldsymbol{\phi},\nu}(\mathfrak{a}) \leq \rho_0^{-4} q^{s(\theta-\nu)} \sum_{\substack{\xi', \eta' \in \mathcal{E}(p^\theta) \\ \xi' \not\equiv \eta' \pmod{p^\nu}}} \rho_\theta(\xi')^2 \rho_\theta(\eta')^2 K_{\theta,\theta,c}^{1,\boldsymbol{\phi}}(\mathfrak{a}; \xi', \eta') = q^{s(\theta-\nu)} K_{\theta,\theta,c}^{1,\boldsymbol{\phi},\nu}(\mathfrak{a}),$$

and the conclusion follows by substituting this into (5.4.5). □

5.5 The iterative process

Let $k \geq 2$, and suppose that Theorem 5.2.1 holds for smaller exponents. In this section, we make use of the inductive hypothesis and provide the key lemmata underlying our iterative process, before completing the proof of the theorem in Section 5.6.

Lemma 5.5.1. *Let $a, b, r \in \mathbb{N}$ with $1 \leq r \leq k-1$ and $\min\{a, b\} \geq \delta\theta$. Suppose that*

$$ra \leq (k-r+1)b \leq B,$$

and set

$$b' = \lceil (k-r+1)b/r \rceil.$$

Then $K^{r,\boldsymbol{\phi},\nu}_{a,b,c}(\mathfrak{a}) \ll q^{tk^2\nu} K^{r,\boldsymbol{\phi},\nu}_{b',b,c}(\mathfrak{a})$.

Proof. We focus on $K^{r,\boldsymbol{\phi},\nu}_{a,b,c}(\mathfrak{a}; \xi, \eta)$, in which we may assume that $p^\gamma \| (\xi - \eta)$ for some $\gamma < \nu$, and write $\xi - \eta = \omega p^\gamma$ with $(\omega, p) = 1$. We introduce

$$B' = (k-r+1)b - ra - (k-r)\gamma,$$

and in the case $B' \leq \nu$, we apply Lemma 5.4.2 as in [71, Lemma 7.1] to obtain $K^{r,\boldsymbol{\phi},\nu}_{a,b,c}(\mathfrak{a}) \ll q^{tk^2\nu} K^{r,\boldsymbol{\phi},\nu}_{b',b,c}(\mathfrak{a})$.

When $B' > \nu$, we consider the solutions counted by $K^{r,\boldsymbol{\phi},\nu}_{a,b,c}(\mathfrak{a}; \xi, \eta)$ and, via the same argument used in [71, Lemma 7.1], deduce that any such solution satisfies

$$(\omega p^\gamma)^{k-r} \sum_{i=1}^{R} (p^a)^l \big(\Psi_l(u_i) - \Psi_l(v_i) \big) \equiv 0 \pmod{p^{(k-r+1)b}} \quad (1 \leq l \leq r),$$

where $\Psi_l(z) = z^l + p^{a-(k-r)\gamma} \Xi_l(z)$ for some $\Xi_l \in \mathbb{Z}[z]$. Our hierarchy (5.4.1) allows us to ensure that ν satisfies

$$k\gamma < k\nu \leq \delta a,$$

and therefore we have

$$a - (k-r)\gamma > (1-\delta)a > \tau B,$$

so the system of polynomials $\boldsymbol{\Psi}$ is p^c-spaced for some $c > \tau(k-r+1)b$, and

satisfies

$$\sum_{i=1}^{R} \Psi_l(u_i) \equiv \sum_{i=1}^{R} \Psi_l(v_i) \pmod{p^{B'}} \quad (1 \le l \le r).$$

Further manipulations, as in [71, Lemma 7.1], lead to the conclusion that

$$K^{r,\boldsymbol{\phi},\nu}_{a,b,c}(\mathfrak{a}) = \oint_{p^B} U^{B'}_{R,r}(\mathfrak{c}) \, |f_b(\boldsymbol{\alpha}, \eta)|^{2s-2R} \, d\boldsymbol{\alpha},$$

where $\mathfrak{c}_u = \mathfrak{a}_{p^a u+\xi} \, e\big(\psi(p^a u + \xi; \boldsymbol{\alpha})\big)$. At this point, we apply the inductive hypothesis, in the form of Corollary 5.2.2, to deduce that

$$U^{B'}_{R,r}(\mathfrak{c}) \ll q^{B'\epsilon^2} U^{B',H'}_{R,r}(\mathfrak{c}).$$

Applying Lemma 5.4.2, using (5.2.5), and rearranging, as in [71, Lemma 7.1], we obtain

$$K^{r,\boldsymbol{\phi},\nu}_{a,b,c}(\mathfrak{a}) \ll q^{B'\epsilon^2 + R(b'-a-H')} K^{r,\boldsymbol{\phi},\nu}_{b',b,c}(\mathfrak{a}).$$

Finally, we have

$$R(b'-a-H') = tr(r+1)(b'-a-H')/2 < tk^2\nu/2,$$

and so, using the hierarchy (5.4.1), we see that

$$K^{r,\boldsymbol{\phi},\nu}_{a,b,c}(\mathfrak{a}) \ll q^{tk^2\nu} K^{r,\boldsymbol{\phi},\nu}_{b',b,c}(\mathfrak{a}),$$

as required. □

From now on we drop any reference to $\boldsymbol{\phi}, \nu$ and c in our notation, since they are assumed to remain fixed. Let $a, b, r \in \mathbb{N}$ satisfy the hypotheses of Lemma 5.5.1, and let $b' = \lceil (k-r+1)b/r \rceil$.

Lemma 5.5.2. *For $r \ge 2$, we have*

$$K^{r}_{a,b}(\mathfrak{a}) \ll q^{tk^2\nu} K^{k-r}_{b,b'}(\mathfrak{a})^{1/(k-r+1)} K^{r-1}_{b',b}(\mathfrak{a})^{(k-r)/(k-r+1)}.$$

When $r = 1$, we have

$$K^{1}_{a,b}(\mathfrak{a}) \ll q^{tk^2\nu} K^{k-1}_{b,kb}(\mathfrak{a})^{1/k} U^{B,b}_{s,k}(\mathfrak{a})^{1-1/k}.$$

Proof. We first note that

$$\begin{aligned}&\frac{tk(k+1)-t(k-r)(k-r+1)}{k-r+1}+\frac{tr(r-1)(k-r)}{k-r+1}\\&=\frac{tr(kr-r^2+k+1)}{k-r+1}=\frac{tr(r+1)(k-r+1)}{k-r+1}=2R,\end{aligned}$$

and therefore

$$\begin{aligned}2s-2R&=\frac{tk(k+1)(k-r)+t(k-r)(k-r+1)}{k-r+1}-\frac{tr(r-1)(k-r)}{k-r+1}\\&=\frac{t(k-r)(k-r+1)}{k-r+1}+\frac{(tk(k+1)-tr(r-1))(k-r)}{k-r+1}.\end{aligned}$$

As such, we may apply Hölder's inequality, as in [71, Lemma 8.1], to obtain

$$K^r_{b',b}(\mathfrak{a};\xi,\eta)=\oint_{p^B}\left|f_{b'}(\boldsymbol{\alpha},\xi)^{2R}f_b(\boldsymbol{\alpha},\eta)^{2s-2R}\right|\,d\boldsymbol{\alpha}\leq U_3^{1/(k-r+1)}U_4^{(k-r)/(k-r+1)},$$

where

$$U_3=\oint_{p^B}\left|f_b(\boldsymbol{\alpha},\eta)^{t(k-r)(k-r+1)}f_{b'}(\boldsymbol{\alpha},\xi)^{tk(k+1)-t(k-r)(k-r+1)}\right|\,d\boldsymbol{\alpha}=K^{k-r}_{b,b'}(\mathfrak{a};\eta,\xi)$$

and

$$U_4=\oint_{p^B}\left|f_{b'}(\boldsymbol{\alpha},\xi)^{tr(r-1)}f_b(\boldsymbol{\alpha},\eta)^{tk(k+1)-tr(r-1)}\right|\,d\boldsymbol{\alpha}=K^{r-1}_{b',b}(\mathfrak{a};\xi,\eta).$$

We may therefore bound $K^r_{b',b}(\mathfrak{a})$ by

$$\rho_0^{-4}\sum_{\xi\in\mathcal{E}(p^{b'})}\sum_{\substack{\eta\in\mathcal{E}(p^b)\\ \xi\not\equiv\eta\pmod{p^\nu}}}\rho_{b'}(\xi)^2\rho_b(\eta)^2K^{k-r}_{b,b'}(\mathfrak{a};\eta,\xi)^{1/(k-r+1)}K^{r-1}_{b',b}(\mathfrak{a};\xi,\eta)^{(k-r)/(k-r+1)},$$

which allows us to see that

$$K^r_{b',b}(\mathfrak{a})\ll K^{k-r}_{b,b'}(\mathfrak{a})^{1/(k-r+1)}K^{r-1}_{b',b}(\mathfrak{a})^{(k-r)/(k-r+1)}$$

by applying Hölder's inequality again, so when $r\geq 2$ we are done by Lemma 5.5.1. When $r=1$, we have $b'=kb$, and therefore

$$K^1_{a,b}(\mathfrak{a})\ll q^{tk^2\nu}K^{k-1}_{b,kb}(\mathfrak{a})^{1/k}K^0_{kb,b}(\mathfrak{a})^{(k-1)/k}.$$

Observing that $K^0_{kb,b}(\mathfrak{a}) = U^{B,b}_{s,k}(\mathfrak{a})$ gives the claimed result. □

We now present a series of lemmata in which we bound the normalised version of our mean values, and we write $\widetilde{K}^r_{a,b}(\mathfrak{a})$ as shorthand for $\widetilde{K}^r_{a,b}(\mathfrak{a})_\Lambda = \widetilde{K}^{r,\boldsymbol{\phi},\nu}_{a,b,c}(\mathfrak{a})_\Lambda$.

Lemma 5.5.3. *Let $b \le (1-\delta)B/k$. When $r \ge 2$, we have*

$$\widetilde{K}^r_{a,b}(\mathfrak{a}) \ll q^{tk^2\nu}\widetilde{K}^{k-r}_{b,b'}(\mathfrak{a})^{1/(k-r+1)}\widetilde{K}^{r-1}_{b',b}(\mathfrak{a})^{1-1/r}.$$

When $r = 1$, we have

$$\widetilde{K}^1_{a,b}(\mathfrak{a}) \ll q^{2tk^2\nu}\widetilde{K}^{k-1}_{b,kb}(\mathfrak{a})^{1/k}(q^{-b})^{\Lambda(1-1/k)}.$$

Proof. As in [71, Lemma 8.2], when $r \ge 2$ we use Lemma 5.5.2 and the definition (5.2.6) to see that

$$\widetilde{K}^r_{a,b}(\mathfrak{a}) \ll (q^{tk^2\nu})^{(k-1)/r(k-r)}\widetilde{K}^{k-r}_{b,b'}(\mathfrak{a})^{1/(k-r+1)}\widetilde{K}^{r-1}_{b',b}(\mathfrak{a})^{(r-1)/r},$$

which leads directly to the desired conclusion since $(k-1)/r(k-r) \le 1$ for $1 \le r \le k-1$. When $r = 1$, we have

$$\widetilde{K}^1_{a,b}(\mathfrak{a}) \ll q^{tk^2\nu}\widetilde{K}^{k-1}_{b,kb}(\mathfrak{a})^{1/k}V^{1-1/k},$$

where

$$V = \frac{U^{B,b}_{s,k}(\mathfrak{a})}{q^{\Lambda H}U^{B,H}_{s,k}(\mathfrak{a})}.$$

By (5.4.3), we have

$$U^{B,b}_{s,k}(\mathfrak{a}) \ll (q^{H-b})^{\Lambda+\epsilon}U^{B,H}_{s,k}(\mathfrak{a}),$$

and consequently

$$V \ll q^{\epsilon H - b(\Lambda+\epsilon)} \ll q^{s\nu-\Lambda b}.$$

Hence we have

$$\begin{aligned}\widetilde{K}^1_{a,b}(\mathfrak{a}) &\ll q^{tk^2\nu}\widetilde{K}^{k-1}_{b,kb}(\mathfrak{a})^{1/k}(q^{s\nu-\Lambda b})^{1-1/k}\\ &\ll q^{2tk^2\nu}\widetilde{K}^{k-1}_{b,kb}(\mathfrak{a})^{1/k}(q^{-\Lambda b})^{1-1/k}\end{aligned}$$

since $s = tk(k+1)/2 \le tk^2$. □

For $1 \le j \le k-1$, we write $\rho_j = j/(k-j+1)$ and $b_j = \lceil b/\rho_j \rceil$.

Lemma 5.5.4. *Let $1 \le r \le k-1$ and $a \ge \delta\theta$ and $b \ge k\delta\theta$ with $ra \le (k-r+1)b$. Then for $kb \le (1-\delta)B$, we have*

$$\widetilde{K}^r_{a,b}(\mathfrak{a}) \ll q^{(r+1)tk^2\nu}(q^{-b})^{\Lambda(1-1/k)/r} \prod_{j=1}^{r} \widetilde{K}^{k-j}_{b,b_j}(\mathfrak{a})^{\rho_j/r}.$$

Proof. When $r = 1$, this follows immediately from Lemma 5.5.3. For $r \ge 2$, we proceed inductively, as in [71, Lemma 9.1]. Suppose that the conclusion is known for all $r < r_0$ for some $2 \le r_0 \le k-1$. By Lemma 5.5.3, we have

$$\widetilde{K}^{r_0}_{a,b}(\mathfrak{a}) \ll q^{tk^2\nu} \widetilde{K}^{k-r_0}_{b,b_0}(\mathfrak{a})^{1/(k-r_0+1)} \widetilde{K}^{r_0-1}_{b_0,b}(\mathfrak{a})^{1-1/r_0} \tag{5.5.1}$$

with $b_0 = b_{r_0} = \lceil (k-r_0+1)b/r_0 \rceil \ge 2b/k > \delta\theta$. We also have

$$(r_0-1)b_0 \le (r_0-1)\big((k-r_0+1)b/r_0 + 1\big) < (k-r_0+2)b.$$

We may therefore use the inductive hypothesis to bound $\widetilde{K}^{r_0-1}_{b_0,b}(\mathfrak{a})$, obtaining

$$\widetilde{K}^{r_0-1}_{b_0,b}(\mathfrak{a}) \ll q^{r_0tk^2\nu}(q^{-b})^{\Lambda(1-1/k)/(r_0-1)} \prod_{j=1}^{r_0-1} \widetilde{K}^{k-j}_{b,b_j}(\mathfrak{a})^{\rho_j/(r_0-1)}.$$

Substituting this into (5.5.1), and writing $\rho_0 = \rho_{r_0} = r_0/(k-r_0+1)$, we see that

$$\begin{aligned}\widetilde{K}^{r_0}_{a,b}(\mathfrak{a}) &\ll q^{tk^2\nu+(r_0-1)tk^2\nu}(q^{-b})^{\Lambda(1-1/k)/r_0} \widetilde{K}^{k-r_0}_{b,b_0}(\mathfrak{a})^{\rho_0/r_0} \prod_{j=1}^{r_0-1} \widetilde{K}^{k-j}_{b,b_j}(\mathfrak{a})^{\rho_j/r_0} \\ &\ll q^{r_0tk^2\nu}(q^{-b})^{\Lambda(1-1/k)/r_0} \prod_{j=1}^{r_0} \widetilde{K}^{k-j}_{b,b_j}(\mathfrak{a})^{\rho_j/r_0},\end{aligned}$$

and so the lemma follows by induction. □

Lemma 5.5.5. *Suppose that all of the hypotheses of Lemma 5.5.4 hold. Then there exists an integer r' with $1 \le r' \le r$ such that*

$$\widetilde{K}^r_{a,b}(\mathfrak{a}) \ll \big(\widetilde{K}^{k-r'}_{b,b_{r'}}(\mathfrak{a})\big)^{\rho_{r'}}(q^{-b})^{\Lambda/(2k)}.$$

Proof. As in [71, Lemma 9.2], we combine the inequality

$$|z_1 \dots z_n| \leq |z_1|^n + \dots + |z_n|^n$$

with Lemma 5.5.4 to obtain

$$\widetilde{K}^r_{a,b}(\mathfrak{a}) \ll q^{(r+1)tk^2\nu}(q^{-b})^{\Lambda(1-1/k)/r} \sum_{j=1}^{r} \widetilde{K}^{k-j}_{b,b_j}(\mathfrak{a})^{\rho_j}.$$

In particular, for some $1 \leq r' \leq r$, we have

$$\widetilde{K}^r_{a,b}(\mathfrak{a}) \ll q^{(r+1)tk^2\nu}(q^{-b})^{\Lambda(1-1/k)/r} \widetilde{K}^{k-r'}_{b,b_{r'}}(\mathfrak{a})^{\rho_{r'}},$$

so it remains to prove that

$$q^{(r+1)tk^2\nu}(q^{-b})^{\Lambda(1-1/k)/r} \leq (q^{-b})^{\Lambda/(2k)}. \tag{5.5.2}$$

We have $(1-1/k)/r \geq 1/k$ for $1 \leq r \leq k-1$, so

$$q^{(r+1)tk^2\nu}(q^{-b})^{\Lambda(1-1/k)/r} \leq q^{(r+1)tk^2\nu}(q^{-b})^{\Lambda/k}.$$

By our assumptions on b and r, and using (5.4.4), we see that

$$b\Lambda/k \geq \delta\theta\Lambda \geq \delta\mu H\Lambda \text{ and } 2tk^3\nu \geq 2(r+1)tk^2\nu,$$

and by (5.4.1) and (5.4.4), we may choose our parameters to ensure that

$$\delta\mu H\Lambda > 2tk^3\lceil 4\epsilon H\Lambda^{-1}\rceil = 2tk^3\nu,$$

so

$$q^{(r+1)tk^2\nu} \leq q^{b\Lambda/(2k)}$$

and (5.5.2) is proved. □

Lemma 5.5.6. *Let $1 \leq r \leq k-1$, and suppose $a \geq \delta\theta$ and $b \geq k^2\delta\theta$ with $ra \leq (k-r+1)b$. Then whenever $k^2b \leq (1-\delta)B$, there exist integers r' with $1 \leq r' \leq k-1$, as well as $a' \geq \delta\theta$ and $b' \geq k^2\delta\theta$ with $r'a' \leq (k-r'+1)b'$, and*

there exists a real number $0 < \rho \leq (1-1/k)^2$ *satisfying*

$$(1+2/k)b \leq b' \leq k^2 b, \qquad b' = \left\lceil \frac{(r'+1)a'}{k-r'} \right\rceil, \qquad \rho b' \geq b,$$

and such that

$$\widetilde{K}^{r}_{a,b}(\mathfrak{a}) \ll \big(\widetilde{K}^{r'}_{a',b'}(\mathfrak{a})\big)^{\rho}(q^{-b})^{\Lambda/(2k)}.$$

Proof. The hypotheses of Lemma 5.5.5 hold, so we may conclude the existence of r_1, with $1 \leq r_1 \leq r$, such that

$$\widetilde{K}^{r}_{a,b}(\mathfrak{a}) \ll \big(\widetilde{K}^{k-r_1}_{b,b_{r_1}}(\mathfrak{a})\big)^{\rho_{r_1}}(q^{-b})^{\Lambda/(2k)}. \tag{5.5.3}$$

We now wish to apply Lemma 5.5.5 a second time, as in [71, Lemma 9.3], to bound $\widetilde{K}^{k-r_1}_{b,b_{r_1}}(\mathfrak{a})$, so we verify that the hypotheses are met in this case. We have $b \geq k^2\delta\theta \geq \delta\theta$ and

$$b_{r_1} = \left\lceil \frac{b(k-r_1+1)}{r_1} \right\rceil > \frac{b}{k} \geq k\delta\theta,$$

as well as $kb_{r_1} \leq k^2 b \leq (1-\delta)B$. Finally,

$$(k-r_1)b \leq (r_1+1)\left\lceil \frac{b(k-r_1+1)}{r_1} \right\rceil = (r_1+1)b_{r_1},$$

so we deduce that there exists r_2 with $1 \leq r_2 \leq k-r_1$ such that

$$\widetilde{K}^{k-r_1}_{b,b_{r_1}}(\mathfrak{a}) \ll \big(\widetilde{K}^{k-r_2}_{b_{r_1},b_{r_2}}(\mathfrak{a})\big)^{\rho_{r_2}}(q^{-b_{r_1}})^{\Lambda/(2k)},$$

where

$$b_{r_2} = \left\lceil \frac{b_{r_1}(k-r_2+1)}{r_2} \right\rceil.$$

Combining this with (5.5.3), we see that

$$\begin{aligned}\widetilde{K}^{r}_{a,b}(\mathfrak{a}) &\ll \left(\big(\widetilde{K}^{k-r_2}_{b_{r_1},b_{r_2}}(\mathfrak{a})\big)^{\rho_{r_2}}(q^{-b_{r_1}})^{\Lambda/(2k)}\right)^{\rho_{r_1}}(q^{-b})^{\Lambda/(2k)} \\ &\ll \big(\widetilde{K}^{k-r_2}_{b_{r_1},b_{r_2}}(\mathfrak{a})\big)^{\rho_{r_1}\rho_{r_2}}(q^{-b})^{\Lambda/(2k)}.\end{aligned}$$

Setting

$$r' = k-r_2, \quad a' = b_{r_1}, \quad b' = b_{r_2}, \quad \text{and} \quad \rho = \rho_{r_1}\rho_{r_2},$$

it remains to show that these satisfy the conditions set out in the statement of this theorem. The various definitions above imply that $1 \leq r' \leq k-1$ and $a' \geq k\delta\theta \geq \delta\theta$, as well as

$$b' = \left\lceil \frac{b_{r_1}(k-r_2+1)}{r_2} \right\rceil = \left\lceil \frac{(r'+1)a'}{k-r'} \right\rceil,$$

which also gives us $b' \leq kb_{r_1} \leq k^2 b$ and $\rho b' \geq \rho b_{r_1}/\rho_{r_2} \geq \rho b/\rho_{r_2}\rho_{r_1} = b$. We observe that

$$\begin{aligned}\rho &= \frac{r_1}{k-r_1+1} \cdot \frac{r_2}{k-r_2+1} \leq \frac{r_1}{k-r_1+1} \cdot \frac{k-r_1}{r_1+1} \\ &= \left(1 - \frac{1}{k-r_1+1}\right)\left(1 - \frac{1}{r_1+1}\right) \leq (1-1/k)^2,\end{aligned}$$

and consequently $b' \geq b(k/(k-1))^2 \geq (1+2/k)b$, since $k^3 \geq (k-1)^2(k+2)$ for all $k \geq 1$. This also implies $b' \geq b \geq k^2\delta\theta$. Finally, we have

$$(k-r'+1)b' = (r_2+1)b' \geq (r_2+1)\frac{b_{r_1}(k-r_2+1)}{r_2} \geq b_{r_1}(k-r_2) = r'a',$$

and the proof of the lemma is complete. □

5.6 Proof of Theorem 5.2.1

Throughout this section, we consider $k \in \mathbb{N}$ and let $s = tk(k+1)/2$. The case $k = 1$ has been handled in Lemma 5.3.1, so we may assume that $k \geq 2$, and that Theorem 5.2.1 is known for smaller exponents. If $\lambda(s,k) \leq 0$, we are done, so we assume that $\lambda(s,k) = \Lambda > 0$ and work towards a contradiction. As in [71, Section 10], we use Lemma 5.4.3 and our hierarchy (5.4.1) to see that

$$\widetilde{K}^1_{\theta,\theta}(\mathfrak{a}) \gg q^{-2s\theta}. \tag{5.6.1}$$

We now set $N = \lceil 16sk/\Lambda \rceil$, and repeatedly apply Lemma 5.5.6 to obtain sequences $(a_n), (b_n), (r_n)$ and (ρ_n) for $0 \leq n \leq N$, satisfying

$$1 \leq r_n \leq k-1, \quad k^2\delta\theta \leq b_n \leq k^{2n+2}\theta, \quad \delta\theta \leq a_n \leq (k-r_n+1)b_n/r_n,$$

and, for $n \geq 1$,

$$0 < \rho_n \leq (1 - 1/k)^2, \quad \rho_n b_n \geq b_{n-1},$$

and such that

$$\widetilde{K}^1_{\theta,\theta}(\mathfrak{a}) \ll \widetilde{K}^{r_n}_{a_n,b_n}(\mathfrak{a})^{\rho_1 \dots \rho_n} (q^{-\Lambda/(2k)})^{nb_0}, \tag{5.6.2}$$

where the empty product $\rho_1 \dots \rho_n$ for $n = 0$ is interpreted as 1. The initial choice of $a_0 = b_0 = \theta$ and $r_0 = \rho_0 = 1$ therefore trivially satisfies (5.6.2). We prove the existence of such sequences by induction, following the same argument used in [71, Section 10].

Using (5.6.2) in the case $n = N$ in conjunction with (5.6.1), and writing $\rho = \rho_1 \dots \rho_N$, we obtain the bound

$$q^{-2s\theta} \ll \widetilde{K}^{r_N}_{a_N,b_N}(\mathfrak{a})^{\rho} (q^{-\Lambda/(2k)})^{N\theta},$$

and Lemma 5.2.4, together with the assumption that $\lambda(s,k) = \Lambda$, gives

$$\widetilde{K}^{r_N}_{a_N,b_N}(\mathfrak{a}) \ll q^{H\epsilon}.$$

By our hierarchy of constants, we may assume that $H\epsilon \leq \theta$, so that

$$q^{-2s\theta} \ll (q^{\rho - N\Lambda/(2k)})^{\theta}. \tag{5.6.3}$$

We now observe that (5.4.4) implies that q^θ is sufficiently large with respect to s, k and Λ, so (5.6.3) can only hold if

$$4s \geq N\Lambda/(2k).$$

The definition of N leads ultimately to the relation

$$\Lambda \leq 8sk/N \leq \Lambda/2,$$

a contradiction to the assumption that $\lambda(s,k) = \Lambda > 0$, and so Theorem 5.2.1 is proved.

5.7 Proof of Theorem 5.1.1

As in Chapter 4, it suffices to prove Theorem 5.1.1 for X a suitably large power of p; a convenient choice here turns out to be $X = p^H$, for H defined as in Section 5.4. We may also assume that we work with a choice of weights satisfying $\mathfrak{a}_x = 0$ for $x \notin \mathcal{E}(X)$.

By Corollary 5.2.2, we find that

$$\oint_{p^B} |f(\boldsymbol{\alpha})|^{2s}\, d\boldsymbol{\alpha} \ll q^{H\epsilon}\rho_0^{-2} \sum_{\xi \in \mathcal{E}(p^H)} \rho_H(\xi)^2 \oint_{p^B} |f_H(\boldsymbol{\alpha}, \xi)|^{2s}\, d\boldsymbol{\alpha}.$$

By (5.2.1) and the Cauchy–Schwarz inequality, we see that

$$\begin{aligned}\rho_H(\xi)^2\, |f_H(\boldsymbol{\alpha}, \xi)|^2 &= \Bigg| \sum_{\substack{x \in \mathcal{E} \\ x \equiv \xi \pmod{p^H}}} \mathfrak{a}_x e(\psi(x; \boldsymbol{\alpha})) \Bigg|^2 \\ &\leq \Bigg(\sum_{\substack{x \in \mathcal{E}(X) \\ x \equiv \xi \pmod{p^H}}} \mathfrak{a}_x^2 \Bigg) \Bigg(\sum_{\substack{x \in \mathcal{E}(X) \\ x \equiv \xi \pmod{p^H}}} 1 \Bigg) = \mathfrak{a}_\xi^2,\end{aligned}$$

and consequently that

$$\oint_{p^B} |f(\boldsymbol{\alpha})|^{2s}\, d\boldsymbol{\alpha} \ll q^{H\epsilon}\rho_0^{-2} \sum_{\xi \in \mathcal{E}(p^H)} \mathfrak{a}_\xi^2 \ll q^{H\epsilon} \ll p^{H\epsilon}.$$

We therefore have

$$\oint |f(\boldsymbol{\alpha})|^{2s}\, d\boldsymbol{\alpha} \leq \oint_{p^B} |f(\boldsymbol{\alpha})|^{2s}\, d\boldsymbol{\alpha} \ll X^\epsilon,$$

and Theorem 5.1.1 is proved.

Bibliography

[1] R. C. Baker. *Diophantine inequalities*. The Clarendon Press, Oxford University Press, New York, 1986.

[2] R. C. Baker. Small fractional parts of polynomials. *Funct. Approx. Comment. Math.* **55** (2016), no. 1, 131–137.

[3] V. Bentkus and F. Götze. Lattice point problems and distribution of values of quadratic forms. *Ann. of Math. (2)* **150** (1999), no. 3, 977–1027.

[4] K. D. Biggs. On the asymptotic formula in Waring's problem with shifts. *J. Number Theory* **189** (2018), 353–379.

[5] K. D. Biggs. Almost equal summands in Waring's problem with shifts. *Monatsh. Math.* **188** (2019), no. 1, 31–35.

[6] K. D. Biggs. Efficient congruencing in ellipsephic sets: the quadratic case, *in preparation*.

[7] K. D. Biggs. Efficient congruencing in ellipsephic sets: the general case, *in preparation*.

[8] B. J. Birch. Forms in many variables. *Proc. Roy. Soc. Ser. A* **265** (1961/1962), 245–263.

[9] J. Bourgain. Fourier transform restriction phenomena for certain lattice subsets and applications to nonlinear evolution equations. I. Schrödinger equations. *Geom. Funct. Anal.* **3** (1993), no. 2, 107–156.

[10] J. Bourgain. On the Vinogradov mean value. *Proceedings of the Steklov Institute of Mathematics* **296** (2017), no. 1, 30–40.

[11] J. Bourgain and C. Demeter. The proof of the l^2 decoupling conjecture. *Ann. of Math. (2)* **182** (2015), no. 1, 351–389.

[12] J. Bourgain, C. Demeter, and L. Guth. Proof of the main conjecture in Vinogradov's mean value theorem for degrees higher than three. *Ann. of Math. (2)* **184** (2016), no. 2, 633–682.

[13] T. D. Browning and S. M. Prendiville. A transference approach to a Roth-type theorem in the squares. *Int. Math. Res. Not. IMRN* **2017** (2017), no. 7, 2219–2248.

[14] S. Chow. Sums of cubes with shifts. *J. Lond. Math. Soc. (2)* **91** (2015), no. 2, 343–366.

[15] S. Chow. Waring's problem with shifts. *Mathematika* **62** (2016), no. 1, 13–46.

[16] S. Chow. Roth–Waring–Goldbach. *Int. Math. Res. Not. IMRN* **2018** (2018), no. 8, 2341–2374.

[17] S. Col. *Propriétés multiplicatives d'entiers soumis à des conditions digitales*. Ph.D. thesis, Université Henri Poincaré, 2006.

[18] D. Daemen. The asymptotic formula for localized solutions in Waring's problem and approximations to Weyl sums. *Bull. Lond. Math. Soc.* **42** (2010), no. 1, 75–82.

[19] D. Daemen. Localized solutions in Waring's problem: the lower bound. *Acta Arith.* **142** (2010), no. 2, 129–143.

[20] H. Davenport. On Waring's problem for fourth powers. *Ann. of Math. (2)* **40** (1939), 731–747.

[21] H. Davenport. *Analytic methods for Diophantine equations and Diophantine inequalities*. Second edition. Cambridge University Press, Cambridge, 2005.

[22] H. Davenport and H. Heilbronn. On indefinite quadratic forms in five variables. *J. London Math. Soc.* **21** (1946), 185–193.

[23] L. E. Dickson. All integers except 23 and 239 are sums of eight cubes. *Bull. Amer. Math. Soc.* **45** (1939), 588–591.

[24] L. E. Dickson. *History of the theory of numbers. Vol. II: Diophantine analysis*. Chelsea Publishing Co., New York, 1966.

[25] P. Erdős, C. Mauduit, and A. Sárközy. On arithmetic properties of integers with missing digits. I. Distribution in residue classes. *J. Number Theory* **70** (1998), no. 2, 99–120.

[26] D. E. Freeman. Asymptotic lower bounds for Diophantine inequalities. *Mathematika* **47** (2000), no. 1-2, 127–159 (2002).

[27] D. E. Freeman. Asymptotic lower bounds and formulas for Diophantine inequalities. In *Number theory for the millennium, II (Urbana, IL, 2000)*, 57–74. A K Peters, Natick, MA, 2002.

[28] B. Green. Roth's theorem in the primes. *Ann. of Math. (2)* **161** (2005), no. 3, 1609–1636.

[29] G. H. Hardy and J. E. Littlewood. A new solution of Waring's Problem. *Quart. J. Math. Oxford* **48** (1920), 272–293.

[30] G. H. Hardy and J. E. Littlewood. Some problems of 'Partitio Numerorum'; I: A new solution of Waring's Problem. *Nachr. Ges. Wiss. Goettingen, Math.-Phys. Kl.* **1920** (1920), 33–54.

[31] G. H. Hardy and J. E. Littlewood. Some problems of 'Partitio numerorum' (VI): Further researches in Waring's Problem. *Math. Z.* **23** (1925), no. 1, 1–37.

[32] G. H. Hardy and S. Ramanujan. Asymptotic Formulæ in Combinatory Analysis. *Proc. London Math. Soc. (2)* **17** (1918), 75–115.

[33] H. A. Helfgott. The ternary Goldbach problem, arXiv:1501.05438v2.

[34] H. A. Helfgott and D. J. Platt. Numerical verification of the ternary Goldbach conjecture up to $8.875 \cdot 10^{30}$. *Exp. Math.* **22** (2013), no. 4, 406–409.

[35] D. Hilbert. Beweis für die Darstellbarkeit der ganzen Zahlen durch eine feste Anzahl n^{ter} Potenzen (Waringsches Problem). *Math. Ann.* **67** (1909), no. 3, 281–300.

[36] L.-K. Hua. On Waring's problem. *The Quarterly Journal of Mathematics* **os-9** (1938), no. 1, 199–202.

[37] L. K. Hua. *Additive theory of prime numbers.* Translations of Mathematical Monographs, Vol. 13. American Mathematical Society, Providence, R.I., 1965.

[38] A. V. Kumchev and T. D. Wooley. On the Waring–Goldbach problem for seventh and higher powers. *Monatsh. Math.* **183** (2017), no. 2, 303–310.

[39] I. Łaba and M. Pramanik. Maximal operators and differentiation theorems for sparse sets. *Duke Math. J.* **158** (2011), no. 3, 347–411.

[40] J.-L. Lagrange. Démonstration d'un théorème d'arithmétique. In *Œuvres de Lagrange, Tome Troisième*, 189–201. Gauthier-Villars (Paris), 1867–1892.

[41] E. Landau. Über die Anzahl der Gitterpunkte in geweissen Bereichen. *Nachr. Ges. Wiss. Goettingen, Math.-Phys. Kl.* **1912** (1912), 687–770.

[42] E. Landau. *Handbuch der Lehre von der Verteilung der Primzahlen. 2 Bände.* Chelsea Publishing Co., New York, 1953. 2d ed, With an appendix by Paul T. Bateman.

[43] U. V. Linnik. On the representation of large numbers as sums of seven cubes. *Rec. Math. [Mat. Sbornik] N. S.* **12(54)** (1943), 218–224.

[44] K. Mahler. Note on Hypothesis K of Hardy and Littlewood. *J. London Math. Soc.* **11** (1936), no. 2, 136–138.

[45] K. Mahler. On the fractional parts of the powers of a rational number. II. *Mathematika* **4** (1957), 122–124.

[46] J. Maynard. Primes with restricted digits. *Invent. Math.* **217** (2019), no. 1, 127–218.

[47] J. Maynard. Primes and polynomials with restricted digits, arXiv:1510.07711.

[48] K. O'Bryant. A complete annotated bibliography of work related to Sidon sequences. *Electron. J. Combin.* (2004), Dynamic Survey 11.

[49] S. T. Parsell and T. D. Wooley. Exceptional sets for Diophantine inequalities. *Int. Math. Res. Not. IMRN* **2014** (2014), no. 14, 3919–3974.

[50] L. B. Pierce. The Vinogradov mean value theorem [after Wooley, and Bourgain, Demeter and Guth]. *Astérisque* (2019), no. 407, Exp. No. 1134, 479–564. Séminaire Bourbaki. Vol. 2016/2017. Exposés 1120–1135.

[51] R. Remak. Bemerkung zu Herrn Stridsbergs Beweis des Waringschen Theorems. *Math. Ann.* **72** (1912), no. 2, 153–156.

[52] K. Roth. Sur quelques ensembles d'entiers. *C. R. Acad. Sci. Paris* **234** (1952), 388–390.

[53] S. Siksek. Every integer greater than 454 is the sum of at most seven positive cubes. *Algebra Number Theory* **10** (2016), no. 10, 2093–2119.

[54] S. B. Stečkin. Mean values of the modulus of a trigonometric sum. *Trudy Mat. Inst. Steklov.* **134** (1975), 283–309, 411. Theory of functions and its applications (Collection of articles dedicated to Sergeĭ Mihaĭlovič Nikol'skiĭ on the occasion of his seventieth birthday).

[55] R. C. Vaughan. On Waring's problem for smaller exponents. II. *Mathematika* **33** (1986), no. 1, 6–22.

[56] R. C. Vaughan. *The Hardy–Littlewood method*. Second edition. Cambridge University Press, Cambridge, 1997.

[57] R. C. Vaughan and T. D. Wooley. Waring's problem: a survey. In *Number theory for the millennium, III (Urbana, IL, 2000)*, 301–340. A K Peters, Natick, MA, 2002.

[58] I. M. Vinogradov. Sur le théorème de Waring. *Bull. Acad. Sci. URSS* **1** (1928), 393–400.

[59] I. M. Vinogradov. *The method of trigonometrical sums in the theory of numbers*. Interscience Publishers, London and New York., 1954. Translated, revised and annotated by K. F. Roth and Anne Davenport.

[60] V. H. Vu. On a refinement of Waring's problem. *Duke Math. J.* **105** (2000), no. 1, 107–134.

[61] E. Waring. *Meditationes algebraicæ*. American Mathematical Society, Providence, RI, 1991. Translated from the Latin, edited and with a foreword by Dennis Weeks, With an appendix by Franz X. Mayer, translated from the German by Weeks.

[62] H. Weyl. Über die Gleichverteilung von Zahlen mod. Eins. *Math. Ann.* **77** (1916), no. 3, 313–352.

[63] A. Wiles. Modular elliptic curves and Fermat's last theorem. *Ann. of Math. (2)* **141** (1995), no. 3, 443–551.

[64] T. D. Wooley. Large improvements in Waring's problem. *Ann. of Math. (2)* **135** (1992), no. 1, 131–164.

[65] T. D. Wooley. On Diophantine inequalities: Freeman's asymptotic formulae. In *Proceedings of the Session in Analytic Number Theory and Diophantine Equations*, volume 360. 2003 .

[66] T. D. Wooley. The asymptotic formula in Waring's problem. *Int. Math. Res. Not. IMRN* **2012** (2012), no. 7, 1485–1504.

[67] T. D. Wooley. Vinogradov's mean value theorem via efficient congruencing. *Ann. of Math. (2)* **175** (2012), no. 3, 1575–1627.

[68] T. D. Wooley. Translation invariance, exponential sums, and Waring's problem. In *Proceedings of the International Congress of Mathematicians—Seoul 2014. Vol. II.* Kyung Moon Sa, Seoul, 2014 505–529.

[69] T. D. Wooley. The cubic case of the main conjecture in Vinogradov's mean value theorem. *Adv. Math.* **294** (2016), 532–561.

[70] T. D. Wooley. Discrete Fourier restriction via efficient congruencing. *Int. Math. Res. Not. IMRN* **2017** (2017), no. 5, 1342–1389.

[71] T. D. Wooley. Nested efficient congruencing and relatives of Vinogradov's mean value theorem. *Proc. London Math. Soc. (3)* **118** (2019), no. 4, 942–1016.

[72] E. M. Wright. The representation of a number as a sum of four 'almost equal' squares. *Q. J. Math.* **8** (1937), 278–279.